REPORT SERIES IN AEROSOL SCIENCE
N:o 228 (2020)

PRODUCTION OF CONDENSIBLE VAPOURS FROM MONOTERPENE OXIDATION

OTSO PERÄKYLÄ

Institute for Atmospheric and Earth System Research / Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

Doctoral thesis

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium P674, Yliopistonkatu 3, on August 7th, 2020, at 2 o'clock in the afternoon.

Helsinki 2020

Author's Address: Institute for Atmospheric and Earth System Research / Physics
P.O. Box 64
FI-00014 University of Helsinki
otso.perakyla@helsinki.fi

Supervisors: Associate Professor Mikael Ehn, Ph.D.
Institute for Atmospheric and Earth System Research / Physics
University of Helsinki

Academician Markku Kulmala, Ph.D.
Institute for Atmospheric and Earth System Research / Physics
University of Helsinki

Reviewers: Professor Miikka Dal Maso, Ph.D.
Aerosol Physics Laboratory
Tampere University of Technology

Docent Tomi Raatikainen, Ph.D.
Finnish Meteorological Institute

Opponent: Professor Jesse Kroll, Ph.D.
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology

ISBN 978-952-7276-37-2 (printed version)
ISSN 0784-3496
Helsinki 2020
Unigrafia Oy

ISBN 978-952-7276-38-9 (pdf version)
http://ethesis.helsinki.fi
Helsinki 2020
Helsingin yliopiston verkkojulkaisut

Acknowledgements

The research for this thesis was carried out at the Institute for Earth System and Atmospheric Research (INAR) of the University of Helsinki, previously part of the Department of Physics. I thank the department heads for providing me the facilities for completing the thesis. I also acknowledge the Vilho, Yrjö and Kalle Väisälä Foundation for funding the bulk of my work. I express my gratitude to Professor Miikka Dal Maso and Docent Tomi Raatikainen for reviewing my thesis. I am also extremely grateful to Professor Jesse Kroll for agreeing to be my opponent. I most warmly thank Academician Markku Kulmala for providing me with an excellent environment to do science, and for having confidence in me ever since I started.

My deepest gratitude goes to Associate Professor Mikael Ehn for his constant support and enthusiasm, and for always being available—not something that can be taken for granted. No matter what the question, you always know the answer, or at the very least, where to look for it. I am extremely thankful for such an inspiring supervisor.

I also want to thank my initial supervisor Matthias Vogt. You gave me a flying start in the world of science, and had great confidence in me from day one. Thank you to Professors Tuukka Petäjä and Veli-Matti Kerminen as well for your support in the beginning of this journey.

I would like to thank all of my coauthors, particularly Yanjun and Pontus. I also want to thank all my colleagues over the years, especially the folks I've shared an office with. I thank Liine, my closest ally over the years, for not only being a great colleague, but for being a very good friend as well. I also want to thank you, Olga and Yanjun, for your invaluable friendship. A special thanks also to Chao, Diego, Fede, George, Heikki, Jenni, Lauriane, Lisa, Matthieu, Nina and Tuija. Doing my PhD has been a much more pleasant experience because of all of you.

Outside the work community, thanks to all my friends and family for listening to me babble about my research, and giving me something else to think about. I am extremely grateful to my family, who have always supported me in my studies and career.

My biggest thanks go to my lovely wife, Jadwiga. You have been beside me throughout the whole dissertation process, and brought balance to my life. My life would be so much duller without you. And finally, I'd like to thank our daughter Sade for making the past year so much fun, even if tiring at the same time.

Otso Johannes Peräkylä
University of Helsinki, 2020

Abstract

Concurrently with greenhouse gases, humankind has been emitting aerosol particles and their precursors into the atmosphere. These solid or liquid particles, tiny enough to float in the air, cause adverse health effects as well as a net cooling effect on the Earth's climate, counteracting part of the warming caused by greenhouse gases. The magnitude of this effect is uncertain, leading to uncertainties in projections of future climate. One of the main causes for the uncertainty is our lacking knowledge of the natural, pre-industrial aerosol particles.

A major source of aerosol particles is the oxidation of volatile organic compounds (VOCs). VOCs are emitted into the atmosphere in large quantities, with biogenic emissions dominating globally over anthropogenic ones. In the atmosphere, VOCs such as monoterpenes, the main group of VOCs emitted by the boreal forests, undergo oxidation reactions, producing vapours of lower volatility. Part of the products condense on pre-existing aerosol particles, or may even form new particles altogether. The conversion of monoterpenes into condensible vapours is the main topic of this thesis.

In this thesis, I aimed to **1)** determine which oxidants are important for monoterpene oxidation in the context of new particle formation, **2)** quantify the volatilities of a group of VOC oxidation products, highly oxygenated organic molecules (HOMs), and **3)** develop new data analysis methods to gain new insights into the formation of condensible vapours. To address these aims, I utilized mass spectrometric methods for measuring VOCs and their oxidation products, in both field and laboratory conditions.

First, we found that oxidation of monoterpenes by the hydroxyl radical was likely very important for the growth of newly formed particles. Our results also suggest that multi-generation oxidation reactions are important. Second, we found that monoterpene-derived HOMs are predominantly of low volatility, though also semi-volatile behavior was observed when the HOMs contained eight or less oxygen atoms. Our estimates for the volatilities lie between earlier parametrizations and recent computations. Finally, we developed a new data analysis method for mass spectrometric measurements, based on a novel factorization technique. Our method efficiently uses the high resolution information in the measured spectra, avoiding many of the time consuming and subjective procedures commonly used. It also allowed us to separate new HOM formation processes that could not be found using traditional methods.

Keywords: VOCs, monoterpenes, HOMs, volatility, SOA, binPMF

Contents

List of publications

This thesis consists of an introductory review, followed by four research articles. In the introductory review, these papers are cited according to their roman numerals.

I **Peräkylä, O.**, Vogt, M., Tikkanen, O.-P., Laurila, T., Kajos, M. K., Rantala, P. A., Patokoski, J., Aalto, J., Yli-Juuti, T., Ehn, M., Sipilä, M., Paasonen, P., Rissanen, M., Nieminen, T., Taipale, R., Keronen, P., Lappalainen, H. K., Ruuskanen, T. M., Rinne, J., Kerminen, V. M., Kulmala, M., Bäck, J., and Petäjä, T. (2014). Monoterpenes' oxidation capacity and rate over a boreal forest: temporal variation and connection to growth of newly formed particles, *Boreal Environ. Res.*, 19:293–310.

II **Peräkylä, O.**, Riva, M., Heikkinen, L., Quéléver, L., Roldin, P., and Ehn, M. (2020). Experimental investigation into the volatilities of highly oxygenated organic molecules (HOMs), *Atmos. Chem. Phys.*, 20:649–669. doi:10.5194/acp-20-649-2020.

III Zhang, Y.*, **Peräkylä, O.***, Yan, C., Heikkinen, L., Äijälä, M., Daellenbach, K. R., Zha, Q., Riva, M., Garmash, O., Junninen, H., Paatero, P., Worsnop, D., and Ehn, M. (2019). A novel approach for simple statistical analysis of high-resolution mass spectra, *Atmos. Meas. Tech.*, 12:3761–3776. doi:10.5194/amt-12-3761-2019. *: **contributed equally.**

IV Zhang, Y., **Peräkylä, O.**, Yan, C., Heikkinen, L., Äijälä, M., Daellenbach, K. R., Zha, Q., Riva, M., Garmash, O., Junninen, H., Paatero, P., Worsnop, D., and Ehn, M. (2020). Insights into atmospheric oxidation processes by performing factor analyses on subranges of mass spectra, *Atmos. Chem. Phys.*, 20:5945–5961. doi:10.5194/acp-20-5945-2020.

Paper I is reprinted by kind permission of Boreal Environment Research. **Papers II**, **III** and **IV** are reprinted under the Creative Commons Attribution 4.0 license.

List of abbreviations

HOM	Highly oxygenated organic molecule
VOC	Volatile organic compound
SVOC	Semi-volatile organic compound
LVOC	Low volatility organic compound
ELVOC	Extremely low volatility organic compound
SOA	Secondary organic aerosol
NPF	New particle formation
CI-APi-TOF	Chemical ionization atmospheric pressure interface time of flight mass spectrometer
PTR-MS	Proton transfer reaction mass spectrometer
UMR	Unit mass resolution
HR	High resolution
PMF	Positive matrix factorization
binPMF	Mass spectral binning combined with PMF

1 Introduction

Over the past few centuries, humankind has become an integral player in shaping the composition of the Earth's atmosphere, changing the Earth's climate and throwing its biogeochemical cycles off balance. The main drivers of these changes have been the continued emissions of greenhouse gases (GHGs), primarily carbon dioxide (CO_2) from fossil fuel combustion. As a result, CO_2 concentrations in the atmosphere have risen from about 280 ppm in pre-industrial times, before 1800, to an annual mean of 410 ppm in 2019 (Etheridge et al., 1996, Ed Dlugokencky and Pieter Tans, NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/, accessed May 12$^{\text{th}}$, 2020)). This is the highest value in hundreds of thousands of years (Lüthi et al., 2008). Going back tens of millions of years, the concentration has been higher, but the rate at which fossil carbon is being released to the atmosphere is unprecedented, at least since the dinosaurs were roaming the Earth 66 million years ago (Zeebe et al., 2016). From pre-industrial times, the global mean temperature has risen by around one degree Celsius (IPCC, 2018).

Greenhouse gases warm the planet through the so-called greenhouse effect. This effect also exists naturally, but increased GHG concentrations are making it stronger, warming the Earth in the process (IPCC, 2013). However, not all human activity has had a warming effect. Concurrently with GHG emissions, mankind has been emitting aerosol particles and their precursors into the atmosphere. These tiny liquid or solid particles, floating in the air, can both warm and cool the planet: their net effect has been a cooling one, with large uncertainties as to its magnitude (IPCC, 2013). This cooling effect has partly counteracted the warming caused by GHGs. Similarly to greenhouse gases, there are both natural and anthropogenic aerosol particles. Important natural aerosol sources include, but are not limited to, sea spray, mineral dust, and volcanic activity. Anthropogenic sources include traffic, industry and biomass burning.

Unlike long-lived greenhouse gases, tropospheric aerosol particles generally only stay in the atmosphere from some days to some weeks. Some are emitted into the atmosphere directly as particles: these are called primary particles, and include e.g. desert dust. Others form in the atmosphere, in the gas-to-particle conversion of low volatility vapours: these are known as secondary particles. Both types originate from both natural and anthropogenic sources. Some aerosol particles are hard to classify: examples include particles from wildfires, exacerbated by anthropogenic activities, and primary particles, on top of which secondary material has condensed.

In general, dark-coloured aerosol particles, such as soot, absorb sunlight and warm the planet. In contrast, light-coloured ones, such as sea spray, scatter sunlight, cooling the planet. In addition to these direct effects, aerosol particles also play a key role in cloud formation: every cloud droplet and ice crystal has been formed around an aerosol particle. Broadly, a higher concentration of aerosol particles results in clouds consisting of more, but smaller cloud droplets: these reflect more sunlight back to space, again cooling the planet (Twomey, 1977). There are also other interactions between clouds and aerosols, and this is an area of active research (Boucher et al., 2013).

Out of all the drivers affecting the climate, the effect of aerosols is by far the most uncertain (IPCC, 2013). This leads to a large uncertainty in the net effect as well. This, for its part, makes it hard to assess how sensitive the climate is to perturbations and makes projections of future climate difficult. The main contributor to the uncertainty of aerosol climate effects is the fact that we do not know what the pre-industrial atmosphere looked like in terms of aerosol particles: thus, assessing how humankind has changed it is difficult (Carslaw et al., 2013). Therefore, to narrow down the uncertainties, understanding the natural state of the atmosphere would be especially important. Before industrialization, the main aerosol sources were mostly natural, with a high contribution from biogenic activity (Andreae, 2007). My work, looking at aerosol formation from biogenic vapours, thus sheds light on aerosol particles not only in the present day, but also in the pre-industrial atmosphere.

In addition to climate effects, aerosol particles affect our lives in other ways as well. They have adverse health effects, being the fifth most important risk factor for premature death and disability (Gakidou et al., 2017). As is evident in polluted megacities, they also degrade visibility. The climate, health and visibility effects of aerosol particles all depend on the size of the particles. The smallest aerosol particles are clusters of just a few molecules, with diameters of less than a nanometre. The upper end of the size range is limited by gravitational deposition: particles larger than some tens of micrometers are sedimented fast enough not to be considered aerosol. Thus, the difference in the diameters of the smallest and the largest particles is over ten-thousandfold. The difference in volume, scaling with the third power of the diameter, is even more mind boggling: a factor of a trillion. Particles of different sizes have very different impacts. The smallest particles do not efficiently interact with visible light: thus, to significantly impact visibility, and to have direct climate effects, an aerosol particle has to be large enough. Sizes of around a hundred nanometres in diameter or more are required. To

act as a cloud seed, somewhat smaller diameters are enough: typically, the minimum required size is around 50 nm (Kerminen et al., 2012). Finally, depending on their size, the fate of particles in the respiratory system can be completely different (Rostami, 2009). As an example, large particles are mainly deposited in the upper respiratory system, while the smallest ones can penetrate deep into the lungs. Thus, the particle size is key in determining its health effects.

Like the size, the chemical composition of aerosol particles varies widely. All across the continental northern hemisphere, a large fraction of the mass of aerosol particles smaller than a micrometre consists of organic material (Q. Zhang et al., 2007). More specifically, the majority of this organic matter is of secondary origin: it has been formed in gas-to-particle conversion of organic vapours (Jimenez et al., 2009). Despite its worldwide significance, the exact formation mechanisms of this secondary organic aerosol (SOA) have long remained elusive (Hallquist et al., 2009). However, considerable progress has been made in recent years (Shrivastava et al., 2017).

SOA is formed when volatile organic compounds (VOCs) get oxidized in the atmosphere, and are transformed into less volatile forms. There are myriads of different VOCs, both anthropogenic and natural ones. Globally, the majority of VOCs are of biogenic origin, with forests being an especially important source (Guenther et al., 2012; Lamarque et al., 2010). In addition to making up a large fraction of the total aerosol mass, organic compounds are also important in new particle formation (NPF) events (Kerminen et al., 2018; Kulmala et al., 1998). In NPF events, low volatility vapours cluster together to form new particles. These particles subsequently grow to larger sizes by the condensation of additional vapours. The vapours taking part in the very first steps of NPF vary from environment to environment. However, their exact identity has remained uncertain. Candidates that have been suggested include e.g. sulfuric acid (Weber & McMurry, 1996), sulfuric acid and ammonia (Kirkby et al., 2011), iodic acid (Sipilä et al., 2016) and organic compounds (Kirkby et al., 2016). As opposed to the uncertainty pertaining the vapours actually forming the very smallest particles, it has been postulated for a long time that the vapours making the particles grow larger, up to CCN sizes, are typically organic (Kerminen et al., 2012; Kulmala et al., 1998; Riipinen et al., 2011; Tunved et al., 2006). Whether NPF events produce climate-active aerosol particles depends mainly on how fast the particles grow, rather than on their formation rate (Kerminen et al., 2012). Hence, organic compounds play a key role in the formation of climate relevant particles from NPF.

Boreal forests emit large quantities of VOCs: among the most emitted compounds are a group called monoterpenes (Guenther et al., 2012; Rinne et al., 2009). Through monoterpenes, boreal forests have been shown to be a potent source of SOA (Tunved et al., 2006). Upon oxidation, monoterpenes form compounds whose volatility spans a vast range (e.g. Donahue et al., 2012). The lower the volatility of the products, the more likely they are to contribute to particle formation. Recently, a previously unknown group of low volatility oxidation products of monoterpenes, and other VOCs, was discovered (Ehn et al., 2014). These compounds, highly oxygenated organic molecules (HOMs), are known to form aerosol efficiently, and even to take part in the very first steps of NPF (Ehn et al., 2014; Kirkby et al., 2016). This implies a low volatility. However, while low, the precise volatilities are still poorly known, hampering our efforts to understand their exact role in aerosol formation (Bianchi et al., 2019).

The discovery of HOMs was made possible by the development of new mass spectrometric measurement techniques. These techniques provide us with an abundance of information about aerosol formation and oxidation chemistry in the atmosphere, which also brings a corresponding amount of data to analyze. The data analysis can be very labour intensive. To further complicate the analysis, many of the choices made during the analysis are subjective, and can impact the end result considerably. As an example, the identification of compounds from the mass spectrum requires skill, and even then it is not guaranteed to produce correct results.

For this thesis, I studied the conversion of the volatile monoterpenes into vapours of low volatility, and how those vapours form aerosol. In addition, I developed novel data analysis techniques to help with the abundance of data produced by mass spectrometric techniques measuring this gas-to-particle conversion. More specifically, I aimed to:

1. Determine monoterpenes' and their oxidants' temporal variation, and their relation to the growth of newly formed particles, at a boreal forest site in southern Finland (**Paper I**).

2. Experimentally characterize the volatilities of highly oxygenated organic molecules (HOMs), and establish what factors affect them (**Paper II**).

3. Develop methods enabling more comprehensive use of the high resolution information in mass spectrometric data, while simplifying the analysis process, and to apply them to ambient data of monoterpene oxidation products in order to gain insights into the oxidation pathways (**Papers III** and **IV**).

2 Formation of secondary organic aerosol from volatile precursors

The formation of SOA starts with the emission of the precursor VOCs into the atmosphere. Their emission sources include both anthropogenic and biogenic ones (Guenther et al., 2012; Lamarque et al., 2010). The formation of aerosol particles from these volatile emissions has been known for a long time: Went (1960) makes the case that the blue haze observed above many forested areas consists of aerosol resulting from the oxidation of volatile organic compounds, citing, among others, studies by John Tyndall almost a hundred years prior (e.g. Tyndall, 1868, 1869). However, the exact molecular mechanisms behind SOA formation remained elusive for another fifty years, and are still not fully understood (Hallquist et al., 2009; Shrivastava et al., 2017).

Globally, the majority of VOCs are from biogenic sources, more specifically from vegetation (Guenther et al., 2012; Lamarque et al., 2010). The reasons that plants emit VOCs are many, and partially unknown, but include defence against herbivores, attraction of pollinators, mechanical damage and even communication between plants (Laothawornkitkul et al., 2009). Most biogenic emissions of VOCs are either light or temperature dependent (Laothawornkitkul et al., 2009). While biogenic sources dominate on a global scale, regionally anthropogenic emissions can be important (Simpson et al., 1999). Anthropogenic VOCs come from sources like traffic, but also from other everyday activities, such as cleaning agents and personal care products, which can be major VOC sources in cities (McDonald et al., 2018).

Many VOCs are highly reactive in the atmosphere, with chemical lifetimes of less than a day (Laothawornkitkul et al., 2009). The main oxidants they react with are ozone (O_3), the hydroxyl (OH) radical, and the nitrate (NO_3) radical. Other oxidants, such as free chlorine (Cl) atoms, can be important in some environments (e.g. Thornton et al., 2010). Upon oxidation, functional groups, such as carbonyls, hydroxyls and carboxyls, are added to the molecule, lowering its volatility due to the ability of these groups to form intermolecular bonds (Jimenez et al., 2009; Ziemann & Atkinson, 2012). It is also possible that the molecule fragments in the process: this increases its volatility (Jimenez et al., 2009; Ziemann & Atkinson, 2012). Depending on the volatility of the product in relation to its concentration in the gas phase, and also on availability of other condensable vapours, the oxidation products may then condense or nucleate to form aerosol (Ziemann & Atkinson, 2012). The higher the concentration, and the lower

the volatility, the more likely aerosol formation becomes. It is also possible that the oxidation products undergo additional chemical reactions, either in the gas or in the particle phase, which may further affect their volatility and propensity for condensation (Ziemann & Atkinson, 2012).

Boreal forests are known to be an important source of SOA (e.g. Tunved et al., 2006). This SOA-forming potential has been attributed to a group of VOCs they emit, called monoterpenes, with the chemical formula $C_{10}H_{16}$. Monoterpenes are among the most abundant VOC emissions from the boreal forests, the mixture of monoterpenes emitted varying from tree to tree and between seasons, with α-pinene being the dominant compound (Bäck et al., 2012; Guenther et al., 2012; Hakola et al., 2012; Rinne et al., 2009). Monoterpene oxidation products are often low-volatile enough to directly contribute to particle formation (e.g. Ehn et al., 2014). They are thought to be important in enhancing the growth of newly formed particles in the boreal forest (e.g. Yli-Juuti et al., 2011), and to be the dominant SOA source in other environments as well, such as in the southeastern United States (H. Zhang et al., 2018). In contrast, the oxidation products of isoprene, the most emitted VOC globally, are more volatile, typically requiring additional particle phase reactions to form SOA in appreciable quantities (Hallquist et al., 2009; Shrivastava et al., 2017). New particle formation events are infrequent in isoprene dominated areas, such as the Amazon rainforest, and isoprene oxidation can even suppress NPF through various mechanisms (Kiendler-Scharr et al., 2009; McFiggans et al., 2019; Wimmer et al., 2018).

For a long time, it was thought that the formation of SOA was mainly driven by the partitioning to the particle phase of semi-volatile organic compounds (SVOCs), that in equilibrium exist in appreciable quantities in both gas- and particle phases (e.g. Donahue et al., 2012; Hallquist et al., 2009; Kroll & Seinfeld, 2008; Seinfeld & Pankow, 2003). This view was motivated by a large fraction of the measured VOC oxidation products being semi-volatile. The view still holds for e.g. the oxidation of isoprene, with the addition that particle phase reactions are important for locking the partitioned mass into the particles (Shrivastava et al., 2017). However, for other precursors such as monoterpenes, it was discovered that observations of SOA formation were best described if a large fraction of the condensing material was assumed to be non-volatile (Barsanti et al., 2011; Pierce et al., 2011; Riipinen et al., 2011). Indeed, this had been suggested already more than a decade earlier by Kulmala et al. (1998), but because of the lack of measurements of suitable condensing vapours, the view didn't catch on.

With the development of new mass spectrometric measurement techniques, such as the atmospheric pressure interface time of flight mass spectrometer (APi-TOF, Junninen et al., 2010), detection of new types of compounds became possible. Among the first observations with the APi-TOF were highly oxygenated organic compounds in the gas phase over the boreal forest (Ehn et al., 2010). These were soon confirmed to be the oxidation products of monoterpenes, including α-pinene, formed directly in the gas phase oxidation, and with presumably low volatilities (Ehn et al., 2012). Using the APi-TOF coupled to a chemical ionization inlet (CI-APi-TOF, Jokinen et al., 2012), the compounds were found to form in large yields in the oxidation of monoterpenes, and to be able to significantly contribute to SOA and new particle formation (Ehn et al., 2014). The CI-APi-TOF is also the main instrument used in this thesis. Based on earlier known relationships between composition and volatility, such as the one presented by Donahue, Epstein, et al. (2011), the compounds were classified as extremely low volatility organic compounds (ELVOCs) (Ehn et al., 2014). This was also supported by the observations that the compounds efficiently condensed to form SOA. Later, it was found that not all of the compounds may be of extremely low volatility (Kurtén et al., 2016). As a consequence, the compounds were renamed as highly oxygenated organic compounds (HOMs): a comprehensive review on what is known of their formation and properties is presented by Bianchi et al. (2019). HOMs have been found to form in a variety of oxidation systems, and to be able to contribute efficiently to the formation of SOA and in NPF (Bianchi et al., 2019; Kirkby et al., 2016; Tröstl et al., 2016). Among the main knowledge gaps on HOMs are the uncertainties in their volatilities and in their fate in the particle phase (Bianchi et al., 2019). Next, I will summarize the theory of VOC oxidation, with an emphasis on HOM formation from monoterpene oxidation.

2.1 Oxidation of VOCs and the formation of HOMs

Volatile organic compounds often have short lifetimes in the atmosphere due to their oxidation reactions. The main oxidants of VOCs are ozone, hydroxyl (OH) radical and nitrate (NO_3) radical. Radicals have an unpaired electron, making them reactive: this electron can be marked in chemical formulae with a dot ($^\bullet$). In the text, I will omit the dot, but in equations I will use it for clarity. Out of the main oxidants, ozone has a relatively low reactivity towards many VOCs, but a significantly higher concentration in the atmosphere. It is also present throughout the day, as opposed to the other two.

OH is mainly a daytime oxidant, formed in photochemistry. In contrast, NO_3 exists mainly in the night, as it is rapidly broken down by daylight and nitric oxide (NO) formed in photochemistry. The oxidants react with VOCs in different ways: ozone and nitrate radical mainly react with double bonds between carbon atoms, if available, while the OH radical can also react by abstracting a hydrogen from the VOC molecule. (Atkinson & Arey, 2003).

HOMs have been observed to form in reactions with of all the main oxidants (Bianchi et al., 2019). In this thesis, I will use the definitions by Bianchi et al. (2019) on what compounds are considered to be HOMs (italicized definitions from Bianchi et al. (2019), explanations by the author):

1. *HOMs are formed via autoxidation involving peroxy radicals.* I will discuss peroxy radicals and autoxidation later in this section. This criterion sets HOMs apart from other organic compounds of high oxygen content, such as sugars.

2. *HOMs are formed in the gas phase under atmospherically relevant conditions.* This condition distinguishes HOMs from compounds formed in e.g. combustion, where autoxidation is known to be important (Cox & Cole, 1985).

3. *HOMs typically contain six or more oxygen atoms.* This last criterion is less strict than the first two. Its main purpose is to emphasize the absolute number of oxygen atoms, as opposed to e.g. the oxygen-to-carbon ratio as a distinguishing feature of HOMs. Some compounds adhering to the first two criteria may be considered to be HOMs even if they only have five oxygen atoms. Correspondingly, some compounds, such as organic nitrates, with six or more oxygen atoms, might not be formed in autoxidation and thus not considered HOMs. However, as there typically is no direct way to deduce whether a compound was formed in autoxidation, the limit of six oxygen atoms is a useful compromise.

In the following discussion, I will focus on the ozone reactions, as these are often the most favourable for HOM formation. Differences and similarities to other oxidants will be highlighted as necessary. I will use cyclohexene as an example VOC. Cyclohexene is a six-carbon ring with a double bond on the ring, making it an endocyclic alkene (Fig. 1). As such, it is a useful surrogate for α-pinene, also containing a six-membered ring with a double bond (Rissanen et al., 2014). In addition, α-pinene has another,

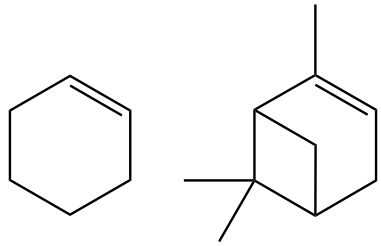

Figure 1: The chemical structures of cyclohexene (left) and α-pinene (right).

four-carbon ring, and three methyl groups, making the molecule more complex (Fig. 1). However, as the initial ozone addition happens on the double bond, cyclohexene provides a convenient simplification. The six-membered ring with a double bond is also found in many other monoterpenes, such as limonene and Δ^3-carene.

The initial reaction between the VOC and the oxidant is only the beginning of an often long cascade of further reactions and intermediates, before relatively stable oxidation products are formed. One important class of intermediates formed in virtually all atmospheric oxidation reactions are organic peroxy (RO_2) radicals (Ziemann & Atkinson, 2012). These consist of an organic structure (the R in RO_2), along with two oxygen atoms bonded to it, forming a chain. The outer oxygen atom has an unpaired electron, making the whole a radical. An example of an RO_2 radical is the $C_6H_9O_4$ on the bottom left in Fig. 2, formed in the ozonolysis of cyclohexene.

In reactions with ozone, the first RO_2 is formed through a couple of steps, outlined here only briefly. First, ozone attaches to the double bond of a VOC, forming a primary ozonide (POZ) (Fig. 2). The POZ rapidly decomposes, with the scission of the bonds between the carbon atoms and within the oxygen ring, forming a carbonyl and a Criegee Intermediate (CI) (Fig. 2; Vereecken & Francisco, 2012). As the bond between the two carbon atoms is broken, the reaction with ozone is also referred to as ozonolysis, with the lysis referring to this cleavage. In compounds where the double bond lies on a ring structure, such as in cyclohexene (Fig. 2) and α-pinene, the scission merely breaks the ring, still keeping the compound in one piece. In contrast, in compounds where the double bond is not on a ring, such as β-pinene, the whole compound is broken in two,

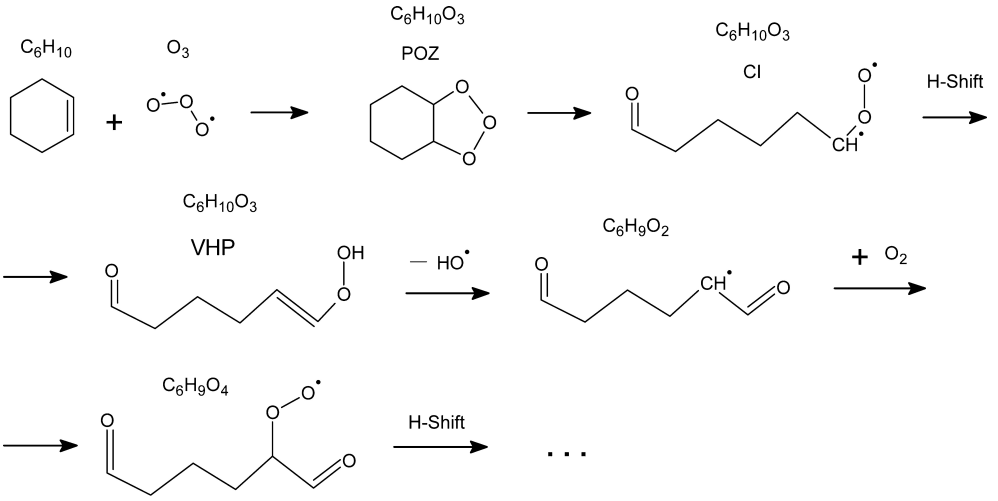

Figure 2: Initial steps of the HOM formation from the ozonolysis of cyclohexene. Note that ozone is unconventionally represented here as a biradical. It is known to exhibit characteristics of both a biradical and a closed shell molecule, with the molecular nature dominating (Miliordos & Xantheas, 2014). Reprinted with permission from Rissanen et al. (2014). Copyright 2014 American Chemical Society.

resulting in products with a carbon number smaller than the parent VOC.

If the CI is stabilized, it can act as an oxidant, forming a significant fraction of the night-time sulfuric acid (Mauldin et al., 2012). However, most often the CI is short lived, and rapidly forms a vinyl hydroperoxide (VHP) (Fig. 2; Vereecken & Francisco, 2012). This can lose an OH radical, itself becoming a carbon centred radical. The carbon centred radical is extremely reactive with atmospheric oxygen, and essentially instantaneously adds an O_2 to form the first peroxy radical ($C_6H_9O_4$ in Fig. 2; Vereecken & Francisco, 2012):

$$R^{\bullet} + O_2 \rightarrow RO_2^{\bullet}. \tag{1}$$

The oxidation of VOCs by OH and NO_3 radicals also forms RO_2 radicals. Both oxidants typically add to the double bond, if one is available (Atkinson & Arey, 2003). Unlike in ozonolysis, this doesn't completely break the bond, but a single bond remains between the carbon atoms. The radical attaches to one carbon, leaving a radical centre on the

other one. Similarly to the ozone reaction, O_2 can then rapidly add, forming an RO_2 radical. Even though all of the oxidants form RO_2 radicals, the radicals are different from each other. In ozonolysis, the double bond is completely broken. The first peroxy radical from the ozone reaction has one less hydrogen as compared to the precursor VOC, and four oxygens (Fig. 2): in the case of cyclohexene (C_6H_{10}), this makes $C_6H_9O_4$, while in the case of α-pinene ($C_{10}H_{16}$), the first RO_2 has the formula $C_{10}H_{15}O_4$. For the OH and NO_3 oxidation as described above, the first peroxy radical consists of the radical added to the precursor VOC, and one added O_2. For α-pinene, this means that the OH-generated RO_2 has the formula $C_{10}H_{17}O_3$, while the NO_3-generated RO_2 has the formula $C_{10}H_{16}O_5N$. Thus, the RO_2 radicals generated by the main oxidants differ from each other in structure, but also in their formulae. It has to be noted, that especially the OH, but also NO_3 radicals, can also abstract a hydrogen from the VOC, changing the composition of the formed RO_2 radicals. However, for compounds with double bonds, this is a minor pathway (Atkinson & Arey, 2003).

In general, RO_2 radicals can undergo a few types of reactions. They can react with radicals, such as nitric oxide (NO), nitrogen dioxide (NO_2), hydroperoxyl (HO_2), and other RO_2 radicals. These reaction channels can either form closed shell products, with no unpaired electrons, or propagate the radical reaction chain. RO_2 can also decompose unimolecularly, with e.g. the loss of an OH radical, leaving a closed shell molecule. I will outline the effect of these reaction pathways later in the text. In addition, in a reaction pathway previously thought to be unimportant in atmospheric chemistry, peroxy radicals can also take part in a process called autoxidation (Crounse et al., 2013). In the process, the peroxy radical centre abstracts a hydrogen atom from elsewhere within the same molecule. In this intramolecular H-abstraction, what was previously a peroxy radical group becomes a hydroperoxide, leaving a carbon centred radical. As in Eq. (1), an oxygen molecule rapidly adds to this radical, resulting in a new RO_2 radical being formed, now with an additional hydroperoxide group attached:

$$RO_2^{\bullet} \xrightarrow{\text{H-shift}} {}^{\bullet}QOOH \xrightarrow{+O_2} HOOQO_2^{\bullet}. \tag{2}$$

Here the ${}^{\bullet}QOOH$ represents the carbon centred radical, where ${}^{\bullet}Q$ has the same structure as the R in $RO_2^{\bullet}$, but with one of the hydrogens abstracted. With a suitable RO_2, the process can repeat multiple times, resulting in the RO_2 gaining high oxygen contents very rapidly (Crounse et al., 2013; Ehn et al., 2014; Jokinen et al., 2014). In the H-shifts followed by autoxidation, one or more O_2 molecules are added to the RO_2 radical.

This conserves the hydrogen and carbon numbers of the initial RO_2 radical, and the parity of the number of oxygen atoms. As an example, the first RO_2 radical generated in the ozonolysis of α-pinene has the formula $C_{10}H_{15}O_4$. If only intramolecular H-shifts and autoxidation act upon the radical, all further RO_2 radicals will also have ten carbon and fifteen hydrogen atoms, and an even number of oxygen atoms. Thus, for an observed RO_2 radical, it is possible to deduce its likely initiating oxidant based on its formula.

Through autoxidation, RO_2 radicals can rapidly gain a high oxygen content. However, most HOM species are not detected as radicals, but closed shell molecules. The transition from a radical to a closed shell molecule happens through a so called termination reaction. As noted above, the RO_2 radical reaction chain can be terminated either uni- or bimolecularly. In unimolecular termination, a smaller radical is split from the RO_2, the main fragment becoming a closed shell product. A common unimolecular termination step is the loss of an OH radical, and the associated formation of a carbonyl: this mechanism can well explain observed spectra of many autoxidation products (Jokinen et al., 2014; Mentel et al., 2015; Rissanen et al., 2014). The loss of a hydroperoxyl radical (HO_2) has also been proposed by e.g. Rissanen et al. (2014), but it seems that this is unlikely for autoxidation products (Hyttinen et al., 2016).

In many environments, the bimolecular reaction of peroxy radicals with NO can dominate. The reaction can proceed *via* two channels, one terminating the radical reaction, the other one propagating it (Vereecken & Francisco, 2012):

$$\mathrm{RO_2^\bullet + NO^\bullet \rightarrow RONO_2} \tag{3a}$$

$$\mathrm{\rightarrow RO^\bullet + NO_2^\bullet}. \tag{3b}$$

In Eq. (3a), a closed shell organic nitrate is formed, while in Eq. (3b), an alkoxy radical is formed. Possible reactions of alkoxy radicals include fragmentation, intramolecular H-shift and reaction with O_2. The first two produce an alkyl radical, enabling a further O_2 addition and formation of another RO_2. It is of note that this alkoxy transition changes the parity of the oxygen number in the radical.

RO_2 radicals can also react with other RO_2 radicals. Again, the possible products include both closed shell and radical species:

$$\mathrm{RO_2^\bullet + R'O_2^\bullet \rightarrow ROOR' + O_2} \tag{4a}$$
$$\rightarrow \mathrm{ROH + R'_{-H}O + O_2} \tag{4b}$$
$$\rightarrow \mathrm{RO^\bullet + R'O^\bullet + O_2}. \tag{4c}$$

In reaction (4a), a covalently bound dimer is formed (Berndt, Mentler, et al., 2018; Berndt, Scholz, et al., 2018; Ehn et al., 2014). These dimers have been shown to be able to take part even in the very first steps of new particle formation, so they possess a special importance for aerosol formation from organic precursors (Ehn et al., 2014; Kirkby et al., 2016; Rose et al., 2018). Dimers were previously thought to be unlikely to form (Vereecken & Francisco, 2012). However, especially for more complex precursors, such as monoterpenes, the formation happens extremely efficiently (Berndt, Scholz, et al., 2018; Ehn et al., 2014). Later studies have found this plausible theoretically as well (Valiev et al., 2019).

Reaction (4b) forms two closed shell monomers. Out of these, ROH has abstracted a hydrogen from the other peroxy radical, forming a hydroxy group, while $\mathrm{R'_{-H}O}$ has lost one, forming a carbonyl. The second product is thus identical with one formed in unimolecular termination by OH loss, while the first one has two more hydrogens. Reaction (4c) forms alkoxy radicals, similarly to Reaction (3b).

Reactions of $\mathrm{RO_2}$ with $\mathrm{HO_2}$ follow a similar pattern as those with other $\mathrm{RO_2}$:

$$\mathrm{RO_2^\bullet + HO_2^\bullet \rightarrow ROOH + O_2} \tag{5a}$$
$$\rightarrow \mathrm{ROH + O_3} \tag{5b}$$
$$\rightarrow \mathrm{RO^\bullet + OH^\bullet + O_2}. \tag{5c}$$

Thus, like the reactions with NO and $\mathrm{RO_2}$, the reactions of $\mathrm{RO_2}$ with $\mathrm{HO_2}$ can either propagate or terminate the radical reaction chain.

Additionally, $\mathrm{RO_2}$ radicals can also react with nitrogen dioxide $\mathrm{NO_2}$:

$$\mathrm{RO_2^\bullet + NO_2^\bullet \rightleftharpoons ROONO_2}. \tag{6}$$

The reaction products are typically unstable, decaying back to RO_2 and NO_2. However, for certain types of RO_2, namely peroxyacyl radicals, and especially at cold temperatures, the products can be long lived, and act as a reservoir species for NO_X (Singh & Hanst, 1981).

In addition to the bimolecular reactions of RO_2 radicals listed here, reactions with other radicals, such as OH and NO_3, are possible as well. However, these radicals occur in the atmosphere in concentrations low enough to make such reactions unlikely.

2.2 HOM formation from different oxidation systems

2.2.1 Laboratory studies

HOMs were initially found in ambient observations, and shown to form in the oxidation of α-pinene (Ehn et al., 2010; Ehn et al., 2012). Later they were quantified by using chemical ionization with nitrate ions (Ehn et al., 2014). Other detection methods, including different ionization chemistries, have also been used to successfully detect HOMs (e.g. Berndt et al., 2016; Berndt et al., 2015; Riva, Rantala, et al., 2019): however, here I will focus mainly on the nitrate CI-APi-TOF results.

The molar HOM yields for α-pinene have been reported to be of the order of five percent (Ehn et al., 2014; Jokinen et al., 2015), with relatively large variation between studies. This variation is caused, for example, by uncertainties in the quantification of HOMs, as well as different experimental setups. Subsequently, HOMs have been shown to form in the oxidation of other types of VOCs as well. Evident examples include other monoterpenes, such as limonene (Ehn et al., 2014; Jokinen et al., 2015). Similarly to α-pinene, limonene contains an endocyclic double bond: in addition, it also has another double bond outside the ring structure, and an even higher HOM yield as compared to α-pinene (Ehn et al., 2014; Jokinen et al., 2015). In addition to the more complex monoterpenes, also cyclohexene, used as an example compound in Sect. 2.1, produces HOMs upon ozonolysis (Ehn et al., 2014; Rissanen et al., 2014). Not all monoterpenes are equally good at forming HOMs: upon ozonolysis, myrcene and β-pinene produce HOMs with yields of less than a percent (Jokinen et al., 2015). These two are either acyclic (myrcene), or have the double bond outside the ring (β-pinene). Thus, it seems that endocyclic compounds are especially favourable for HOM formation. Isoprene has

an even smaller HOM yield, while the sesquiterpene ($C_{15}H_{24}$) β-caryophyllene has a HOM yield of a few percent (Jokinen et al., 2015; Jokinen et al., 2016).

For the aforementioned compounds showing high HOM yields, the yields from ozonolysis are typically higher than from the OH oxidation (Jokinen et al., 2015). Some of this has been attributed to the less efficient detection of OH-produced RO_2: still, the yields remain on the low end, and the products are less oxygenated as compared to ozonolysis reactions (Berndt et al., 2016). All of the compounds listed above contain one or multiple double bonds between carbon atoms, and are thus alkenes. HOM formation has also been reported from the OH oxidation of various aromatics (Garmash et al., 2020; Molteni et al., 2018). Garmash et al. (2020) reported that repeated OH oxidation was important for HOM formation from benzene. Alkenes typically produce considerable HOM yields already from a single oxidation step (Bianchi et al., 2019; Ehn et al., 2014). Even so, the formation of HOM-like compounds has been also reported from the repeated OH oxidation of α-pinene, and the OH oxidation of pinanediol, a representative first-generation oxidation product of α-pinene (Ehn et al., 2014; Schobesberger et al., 2013). As the HOM yields for first-generation oxidation are typically less than 10 % (Bianchi et al., 2019), the majority of the oxidation products are not HOMs, but can potentially become HOMs upon further oxidation.

2.2.2 Ambient observations

In addition to laboratory studies, HOMs have also been observed in the ambient atmosphere: it was, after all, in the boreal forest air that they were first detected. An important study by Yan et al. (2016) utilized positive matrix factorization (PMF, see Sect. 3.2) on measurements of HOMs in the gas phase above the boreal forest. They were able to associate the measured HOMs with different formation pathways, consistent with the RO_2 reactions presented in Sect. 2.1. They found that, during the night-time, the majority of the HOMs were formed in the ozonolysis of monoterpenes, with the RO_2 reactions terminated by other RO_2 according to Reactions (4a) – (4b). Another important contributor to night-time HOM formation was the oxidation of monoterpenes by NO_3 radicals. In the daytime, the majority of the HOMs were found to be associated with reactions with NO. A similar study was conducted by Massoli et al. (2018), but on measurements in Alabama, US. They found broadly similar results, but with high contributions from isoprene oxidation as well.

2.3 Role of HOMs in aerosol formation

HOMs have been experimentally shown to be able to efficiently form SOA (Ehn et al., 2014). In addition to condensing on pre-existing aerosol particles, they have also been shown to be able to take part in the very first steps of new particle formation (Kirkby et al., 2016), and in the early growth of the formed particles (Tröstl et al., 2016). Their importance for particle formation has also been demonstrated in modelling studies (Jokinen et al., 2015; Roldin et al., 2019).

For a compound to take part in the early growth of newly formed particles, it needs to be of low volatility (Donahue et al., 2012; Donahue, Trump, et al., 2011; Kulmala et al., 1998; Riipinen et al., 2011). The exact volatilities of HOMs are poorly known: different estimates of them differ as much as ten orders of magnitude (Kurtén et al., 2016). As a result, assessing role of HOMs in particle formation remains challenging (Bianchi et al., 2019). To address this knowledge gap, I studied the volatilities of HOMs in **Paper II**.

3 Methods

Volatile organic compounds and their oxidation products can be measured using various methods. A major division is that between online and offline methods: in the former, air is continuously sampled, and measured more or less instantaneously, while in the latter, sample air is collected, and then separately analyzed. As a result, online methods generally have a higher time resolution. Benefits of offline methods include an often more detailed chemical characterization of the compounds analyzed. However, offline methods can potentially perturb the sample more than online methods do. As a result, reactive compounds originally present in the air, such as the hydroperoxide-containing HOMs, may have undergone different transformations by the time they are analyzed.

In this work, I have utilized various online measurement methods, most notably mass spectrometry. Broadly, in mass spectrometry, the mass-to-charge ratio of a compound, as well as its abundance in the sample, is determined. Mass spectrometry is well suited for the detection of VOCs at the minute levels they exist in the Earth's atmosphere (Lindinger et al., 1998). Recent developments in mass spectrometry have allowed both the detection of even lower concentrations of VOCs (Breitenlechner et al., 2017; Krechmer et al., 2018), and that of a wide suite of their oxidation products (Riva, Rantala, et al., 2019). The latter is especially important, as the detection of HOMs only became possible through these new developments. I will summarize mass spectrometric methods, and their use in the study of the oxidation of volatile organic compounds in Sect. 3.1.

Mass spectrometry, being able to detect compounds never before detected, also produces a wealth of data to analyze. The analysis is neither straightforward nor trivial, and can be prone to many errors (e.g. Cubison & Jimenez, 2015). In general, the attribution of signal to measured compounds with their mass-to-charge ratios close to each other becomes increasingly difficult as the ratios come closer together. In addition, the identification of compounds based on their mass-to-charge ratio can be challenging and subjective. There are many ways to analyze mass spectrometer data: I will introduce some of these in Sect. 3.2, and outline some challenges in current analysis methods.

3.1 Mass spectrometry

In mass spectrometry, the mass-to-charge ratio (m/z) of a compound is measured. As a result, the measured compounds need to have a charge. Natural ions can be directly measured, while neutral compounds require charging. HOMs were initially detected naturally charged (Ehn et al., 2010), clustered with the nitrate ion (NO_3^-), using the atmospheric pressure interface time of flight mass spectrometer (APi-TOF, Junninen et al., 2010). To measure neutral compounds with the APi-TOF, a chemical ionization (CI) inlet can be coupled to the it (Jokinen et al., 2012). In chemical ionization, the neutral sample molecules are actively charged: multiple different ions can be used, with nitrate ions commonly used for HOM detection. Such a CI-APi-TOF was used by Ehn et al. (2014) to quantify HOMs.

The nitrate ionization is selective for strong acids, such as sulfuric acid, and highly oxygenated compounds (Eisele & Tanner, 1993; Hyttinen et al., 2015). Thus, VOCs themselves, and the majority of their oxidation products, cannot be measured with nitrate ionization. Instead, other techniques are used: probably the most widespread is proton transfer reaction mass spectrometry, where hydronium ions (H_3O^+) protonate VOCs (Lindinger et al., 1998). An instrument utilizing this ionization scheme is called a proton transfer reaction mass spectrometer (PTR-MS). Major developments of the PTR-MS have recently enabled the detection of much lower concentrations of VOCs, as well as more oxygenated products, using instruments such as the PTR3 (Breitenlechner et al., 2017) and VOCUS (Krechmer et al., 2018). In addition to nitrate and PTR, many other ionization schemes exist as well, each with its own strength and weaknesses. Examples include iodide and amine adduct ionization: iodide is commonly used for the detection of moderately oxgenated compounds, while amine can detect many of the same HOMs that are seen with nitrate ionization as well (Riva, Rantala, et al., 2019).

Compounds that share the same molecular formula, but differ in structure, cannot be separated using mass spectrometry without additional analysis steps. As an example, all monoterpenes share the formula $C_{10}H_{16}$. In their protonated form, they are detected as $C_{10}H_{17}^+$, without more detailed information on their identity or chemical structure.

How precisely the m/z of a compound can be determined depends on e.g. the mass resolution of the mass spectrometer, defined as the ratio of the location of a peak to its width. The narrower the peak caused by a compound is, the higher the resolution (Fig. 3). For a mass spectrometer of low resolution, ions situated close to each other

on the mass axis cannot be separated. However, this is not an issue if the ions of interest are the dominant ones at their mass. Cases such as those presented in Fig. 3 are all commonly encountered, including in this thesis. In **Paper I**, the monoterpene measurements were conducted with a quadrupole mass spectrometer, with a resolution of less than 1200. In **Papers III** and **IV**, the field measurements were conducted with an instrument having a resolution of around 3500. Finally, in **Paper II**, the laboratory measurements were done with a resolution of more than 13000.

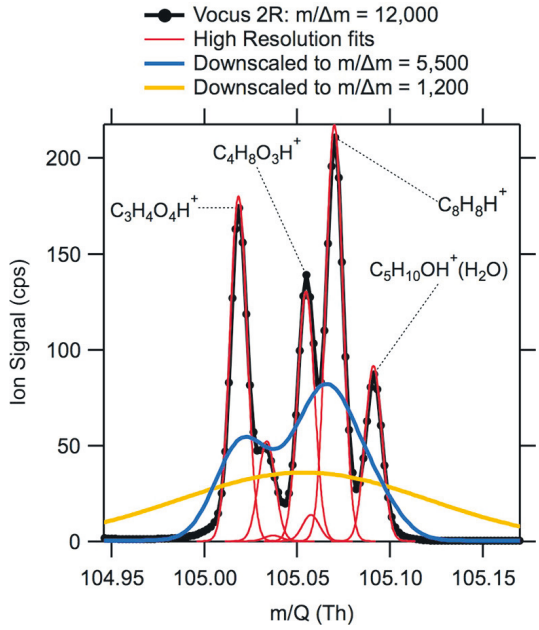

Figure 3: Example mass spectrum showing the effect of mass resolution on the identification of compounds. At the unit mass 105 Th shown here, there are at least six distinct ions. These are well separated at a mass resolution of 12000. At a resolution of 5500, two peaks are clearly distinct, with indications of multiple more. At a resolution of 1200, almost all information of the distinct ions is lost. Reprinted with permission from Krechmer et al. (2018). Copyright 2018 American Chemical Society.

3.2 Data analysis

3.2.1 Mass spectrometry data preprocessing

The parameters we are typically interested in with mass spectrometric measurements are the abundance and time behaviour of a measured compound, such as monoterpenes or their oxidation products. This requires processing of the measured data. Typically, a mass spectrometer records either a full mass spectrum in some defined range, or parts of it, at a certain time resolution. Depending on the type of mass analyzer used, the spectrum first needs to be converted to units of mass-to-charge ratio. With time of flight analyzers, such as the one used in the APi-TOF, the measured quantity is the flight time of ions, which is proportional to the mass-to-charge ratio. This conversion, called mass calibration, is never perfect, and can cause errors further downstream in the data analysis, as the spectrum is not fully aligned with the mass axis.

Figure 3 shows an example of a portion of a spectrum, measured at a single time point, and already mass calibrated. To obtain the time series of a compound, for example $C_3H_4O_4H^+$ in the figure, we need to determine its contribution to the signal at each time point. This can be done using high resolution (HR) peak fitting: with this approach, peaks of pre-defined locations, and often pre-defined widths, are fitted to the observed spectrum using e.g. least squares estimation. Examples of high resolution fits are shown in Fig. 3 in red. This gives us the amplitude of the signal at the given time point, which can often be related to the concentration of the measured species in the sample. Detailed discussion of the calibration techniques required for this are outside the scope of this thesis. High resolution fits to subsequent spectra give us the time behaviour of the signal.

There are multiple challenges with the above described high resolution fitting procedure. First, if the mass resolution of the mass spectrometer is too low, compounds cannot be easily separated from each other (as is the case for resolution 1200 in Fig. 3). Even with a high mass resolution, some compounds can be too close to be unambiguously resolved (e.g. Cubison & Jimenez, 2015). Not only resolution affects this separation: precise mass calibration is required as well. If the spectrum in Fig. 3 would be shifted to the left even by 0.01 Th, the peak fits would be completely thrown off: often this is the leading source of uncertainty in the fits (Cubison & Jimenez, 2015). High resolution fitting also typically requires a pre-defined list of compounds or masses

to fit peaks to. Identification of peaks from a mass spectrum can be subjective and challenging, and in some cases outright impossible (Stark et al., 2015). This is especially true for measurements of the ambient atmosphere. Even with a well constructed peak list, there may be cases where additional, unaccounted for ions are present in the spectrum: in these cases, it is possible that their signal is incorrectly attributed to compounds in the peak list.

To address the problem of uncertain peak assignments, methods have been developed that require no precise identification of the peaks, or even no peak fitting at all, but still use the high resolution information in the spectra (Stark et al., 2015). A crude approach is to sum all the signal at a given integer mass into one lump for each time point: this is often called unit mass resolution (UMR) analysis. This requires no prior knowledge of the ions present in the spectrum, and is computationally easy. However, any high resolution information is lost in the process. In cases where there generally are only individual compounds at each mass this is acceptable, and can even offer a better alternative to high resolution fitting, as the error prone fitting procedure is completely omitted. Also, if the instrumental mass resolution is poor, this is often the only choice.

3.2.2 Finding the sources of compounds: positive matrix factorization

Whether HR or UMR analysis is used, there are often tens or hundreds of variables produced by the technique. As an example, HOMs from α-pinene ozonolysis are typically spread over masses from around 300 Th to around 620 Th, each with their own time behaviour. Analyzing the time series of each of these, and how each is related to the others, can be very time consuming. If there are complex interactions between the variables, these can be nearly impossible to identify from just visual inspection of the time series. In addition, in UMR analysis, one variable can contain contributions from multiple ions. Indeed, the same can be true for HR analysis. And even if the individual ions are perfectly separated in HR fitting, many isomers can exist for ions having the same formula, each having potentially different sources.

In various environments, out of the hundreds of compounds in the HOM mass range, many come from the same source: monoterpene oxidation (Massoli et al., 2018; Yan et al., 2016). Thus, while there is a large number of individual compounds or variables, they all represent one source. In more detail, some of the compounds come from the termination of RO_2 radicals by NO, while some come from termination by other

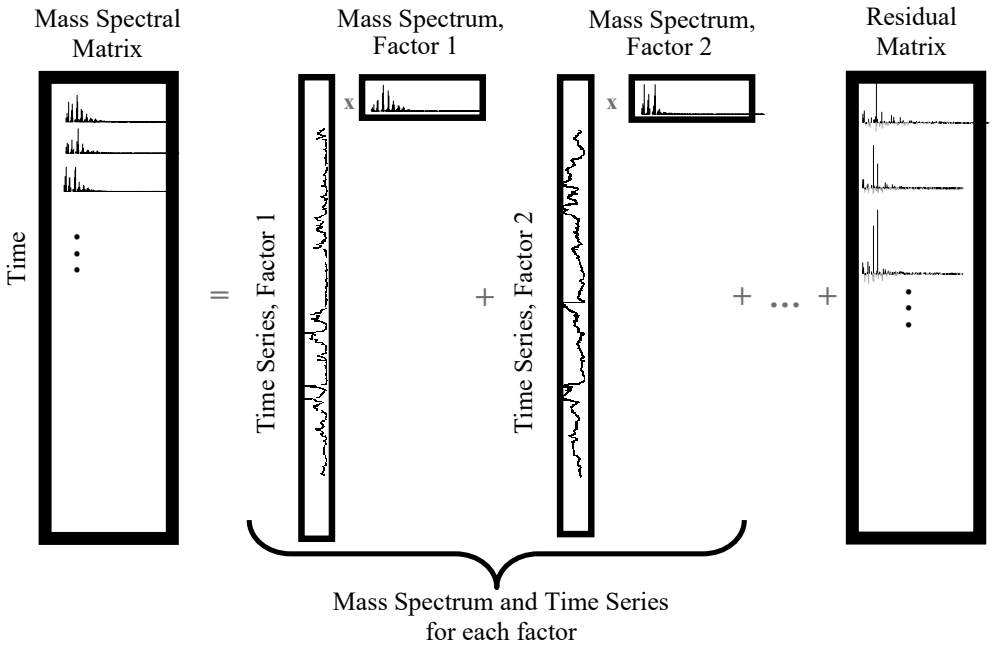

Figure 4: Schematic representation of PMF. The mass spectral matrix, on the left, represents the measured data, and consists of mass spectra measured at multiple time points. This is used as an input to PMF, which breaks the measured spectra down into components, or factors. Each of the factors gets its own time behaviour and mass spectrum, shown on the right hand side. Adding the contributions of the factors together results in an approximation of the measured data. This factorization is never perfect: the difference between the original data and the factorization result is represented by the residual matrix. In finding the PMF solution, this residual is minimized. The factors could represent, for example, HOMs formed in the night-time, upon RO_2 termination, and in the daytime, upon NO termination. Then the spectrum of the first would contain e.g. HOM dimers, and its time series would have high values during the night. The spectrum of the other factor would contain e.g. organic nitrates, and its time series would have high values during the daytime. Figure adopted from Ulbrich et al. (2009) under the Creative Commons Attribution 3.0 License.

RO_2 radicals. NO is strongly light-dependent, with a daytime maximum and very low concentrations during the night. As a result, the majority of the HOMs observed in the daytime can arise from termination by NO (Yan et al., 2016). In contrast, during the night-time, termination by RO_2 may dominate (Yan et al., 2016). While both the HOMs terminated by NO and those by RO_2 depend on monoterpene concentrations, they will have time behaviours very different from each other. In other words, the HOMs from each source will vary together. To uncover such common sources for multiple variables, various statistical techniques can be used. One technique widely used for mass spectrometric data analysis is the positive matrix factorization (PMF, Paatero & Tapper, 1994). PMF tries to reduce the data set, containing potentially hundreds of variables, into a typically much smaller number of factors. Each of these factors has a unique mass spectral profile and time behaviour (Fig. 4). If the factorization is successful, the mass spectrum measured at any given time point can be represented as a sum of the contributions from multiple factors. For example, a daytime HOM spectrum above the boreal forest may contain a high contribution from HOMs formed in RO_2 + NO termination, while the night-time spectrum may be dominated by RO_2 + RO_2 termination. The typical RO_2 + NO termination spectrum is represented by the factor profile, while the contribution of RO_2 + NO termination at any given time point is given by the factor time series: both of these are found by PMF (Fig. 4). PMF has been succesfully used with CI-APi-TOF data to identify different oxidation pathways forming HOMs both with UMR data (Yan et al., 2016) and with HR data (Massoli et al., 2018).

4 Oxidation of monoterpenes above a boreal forest

In **Paper I**, we utilized the long-term measurements at the Hyytiälä SMEAR II station (Station for measuring ecosystem-atmosphere relations, Hari & Kulmala, 2005) to determine the annual and daily cycles of the main oxidants, as well as the effect of monoterpene oxidation on the growth of newly formed particles. The station is located within the boreal forest in southern Finland, at an elevation of 181 m a.s.l., with few sources of anthropogenic pollution nearby: the largest major settlement is the city of Tampere around 50 km southwest. Continuous measurements at the station were started already in 1996, and include various aerosol properties and trace gases, as well as ecosystem functions. In addition, since 2006, volatile organic compounds have been measured at the station with a PTR-MS. Our aim was to determine the relative contributions of the different oxidants to monoterpene oxidation, as well as how this varies in time. With this information, and the measured monoterpene concentrations, we aimed to study how the oxidation of monoterpenes by different oxidants affects the growth of newly formed particles at the site.

Out of the main oxidants, only ozone is continuously measured at the station. Nitrate and hydroxyl radicals are not continuously measured, so we calculated estimates for them using proxies. The OH radical concentration correlates strongly with UVB radiation intensity (Röhrer & Berresheim, 2006), which is measured at the station. Petäjä et al. (2009) presented measurements of sulfuric acid and OH concentrations in Hyytiälä in the spring, and different proxies for sulfuric acid. Utilizing these, we used UVB radiation as a proxy for the OH radical concentration:

$$[\mathrm{OH}]_{\mathrm{proxy}} = 5.62 \times 10^5 \times \mathrm{UVB}^{0.62}, \tag{7}$$

where UVB represents UVB intensity in Wm^{-2} and the concentration of OH is expressed in radicals per cubic centimetre.

The calculation of the NO_3 radical concentration was not as straightforward. We utilized a steady state assumption in the calculation, assuming that on the timescales considered, the production and loss of NO_3 radical are equal. With this assumption, the concentration of the nitrate radical in the air can be expressed as a product of its production rate and its atmospheric lifetime:

$$[\mathrm{NO_3}] = Q_{\mathrm{NO_3}} \tau_{\mathrm{NO_3}}, \tag{8}$$

where $Q_{\mathrm{NO_3}}$ is the production rate of NO_3, and $\tau_{\mathrm{NO_3}}$ is its lifetime. I use here a different symbol for the production rate (Q, as opposed to J used in **Paper I**) for consistency with **Paper II**, where Q was used. NO_3 is produced in the reaction of NO_2 with ozone, both of which are measured: thus, the production rate is straightforward to calculate. For the lifetime, we took into account the reactions of NO_3 with monoterpenes, isoprene and NO, and the equilibrium reaction with NO_2. During the daytime, NO_3 reacts with photochemically produced NO, and is photolyzed rapidly: during these times, we set the lifetime to 5 seconds.

Allan et al. (2000) measured the concentration of the NO_3 radical in the remote marine boundary layer, and also calculated its lifetime and concentration in a similar manner as in **Paper I**. They found that when the lifetime was short, NO_3 was in a steady state and the measured and calculated concentrations agreed well. Measurements of NO_3 have also been attempted in Hyytiälä, but the concentration has always been below detection limit (Liebmann et al., 2018; Rinne et al., 2012). Liebmann et al. (2018) also measured the reactivity of NO_3 towards VOCs: they found that the reactivity, and thus lifetime, is rather well described by reactions with monoterpenes. This validates the steady state calculation approach. In contrast, the much simpler OH estimate (Eq. (7)) has been found to disagree with the modelled OH concentration, and also OH concentration measurements in the summertime (Chen et al., 2020). As a result, the OH concentration is more uncertain.

We then calculated the oxidation capacity of the atmosphere with respect to monoterpenes, as the sum of the concentrations of the different oxidants, weighted with their reaction rate coefficients with monoterpenes:

$$\mathrm{OCAP_{MT}} = k_{\mathrm{OH\ +\ MT}}[\mathrm{OH}] + k_{\mathrm{O_3\ +\ MT}}[\mathrm{O_3}] + k_{\mathrm{NO_3\ +\ MT}}[\mathrm{NO_3}], \tag{9}$$

where $\mathrm{OCAP_{MT}}$ stands for oxidation capacity with respect to monoterpenes. The reaction rate coefficients (k) vary between individual monoterpenes. Further, the mixture of monoterpenes in Hyytiälä varies across the year. To account for this, we calculated the average reaction rate coefficients for each month separately based on the measurements of individual monoterpenes by Hakola et al. (2012). Using the oxidation

capacity and the measured monoterpene concentrations, we could then calculate the monoterpene oxidation rate:

$$\mathrm{OR_{MT}} = \mathrm{OCAP_{MT}}[\mathrm{MT}]. \tag{10}$$

Our motivation for the study was to determine how the oxidation of monoterpenes affects the growth of newly formed particles in the boreal forest. To this end, we calculated the growth rates of newly formed particles in three size classes, 1.5 – 3 nm, 3 – 7 nm and 7 – 20 nm, using the method presented by Hirsikko et al. (2005) and Yli-Juuti et al. (2011).

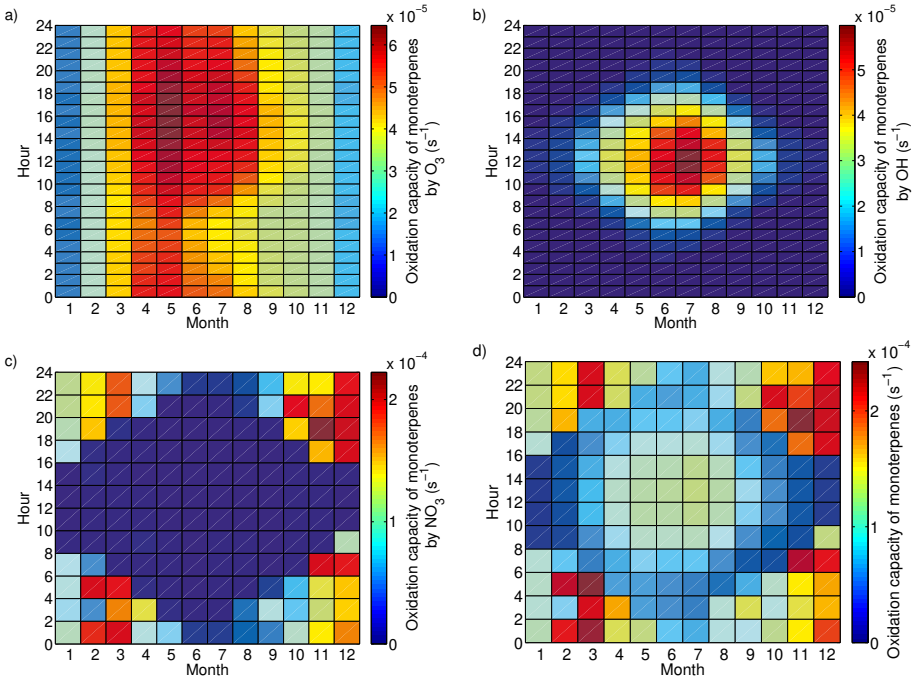

Figure 5: Temporal variation in median oxidation capacity of monoterpenes, by **(a)** O_3, **(b)** OH, **(c)** NO_3 and **(d)** all oxidants. In figures including the NO_3 concentration, the averaging is over two-hour interval. Note the differing colour scales. Figure from **Paper I**.

We found that the proxy OH and nitrate radicals had prominent and opposite diurnal and annual cycles, as expected: OH was most abundant in the daytime and in the

summer, while NO_3 was mainly a night- and wintertime oxidant (Fig. 5). Ozone had less steep, but still prominent, annual and diurnal cycles, with a maximum in the springtime afternoon (Fig. 5). The highest absolute values, typically around 2×10^{-4} s^{-1}, were mostly found in the autumn and springtime nights, when the oxidation capacity by NO_3 was the highest. This corresponds to a monoterpene lifetime of a bit more than an hour. In terms of monoterpene oxidation, however, these winter- and nighttime peaks are relatively uninteresting, as monoterpene emissions are low during those times.

To elucidate the impact of monoterpene oxidation on nanoparticle growth, we plotted the growth rates of particles of different size classes during NPF events against multiple independent variables. The independent variables we investigated were monoterpene concentrations, the oxidation rates of monoterpenes by the individual oxidants, as well as the total oxidation rate, and the oxidation capacities, again both by individual oxidants and as a total. We calculated these parameters for various time windows: during the particle growth in the size class in question, between sunrise and the beginning of the growth, and during the night preceding the event.

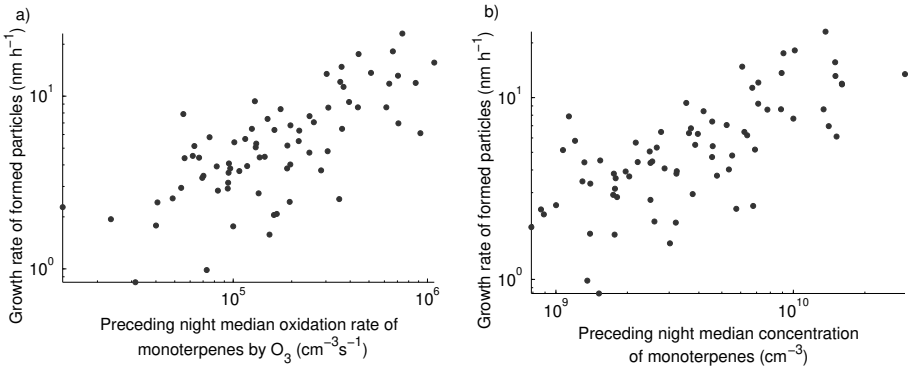

Figure 6: Growth rates of 7 – 20 nm particles, plotted against (a) median oxidation rate of monoterpenes by O_3 during the preceding night, and (b) the median concentration of monoterpenes during the preceding night. Figure from **Paper I**.

We did not find strong correlations of any of the investigated parameters with the growth of the smallest particles, in size classes 1.5 – 3 nm and 3 – 7 nm. Thus, from now on I will refer to the growth rate of 7 – 20 nm particles simply as the growth rate. Out of the quantities measured during the growth of the particles, the highest correlation with the growth rate was found for the oxidation rate of monoterpenes with

OH (both log-transformed, $r = 0.64, p = 3.3 \times 10^{-8}, n = 60$). This points to the OH oxidation during the event contributing to particle growth. However, we found an even stronger correlation of the growth rate with the monoterpene concentrations, and their oxidation rate by O_3, measured in the night preceding the NPF event (again both log-transformed, $r = 0.71, p = 2.0 \times 10^{-12}, n = 74$ and $r = 0.72, p = 1.1 \times 10^{-12}, n = 73$, respectively, Fig. 6).

It is curious that the highest correlation with the growth rate was found between the ozone oxidation rate of monoterpenes during the preceding night. Any low volatility vapours formed during the night will condense long before the start of the NPF event on the following day, and thus be unable to contribute to the growth during the event. Thus, if there is a causal link, it would have to be an indirect one. The other possibility is that there is some confounder or other artefact explaining the correlation. Further, between the night and the event, the planetary boundary layer grows in height, leading to a dilution that is variable from day to day. Such variable dilution should further weaken the connection of the preceding night and the event of the following day.

In **Paper I**, we hypothesized that there may be a nightly build-up of first-generation oxidation products of monoterpenes, that are too volatile to efficiently condense on particles. These products would be formed in the ozonolysis of monoterpenes, upon which they lose a carbon-carbon double bond. As most monoterpenes only have one such bond, the reaction products would be unreactive towards ozone, and only slightly reactive towards NO_3. However, they could still react with OH through hydrogen abstraction. These compounds, lacking an efficient removal mechanism, would build up in the air during the night, and could then be oxidized by OH after the sunrise. This further oxidation might lower their volatility enough to enable their condensation on the newly formed particles. As products formed in the OH-oxidation of first generation monoterpene oxidation products have been shown to be able to take part in new particle formation (Schobesberger et al., 2013), this remains a viable hypothesis.

No further proof of this hypothesis has been presented, and thus the possibility of the link being an artefact remains. One way to further investigate the potential link would be through the measurements of HOMs. These are observed in Hyytiälä in concentrations high enough to explain the observed growth rates of the newly formed particles (Ehn et al., 2014): thus, investigating their sources could shine a light on the causes of the correlation between the preceding night and the NPF event.

5 Volatilities of HOMs

In **Paper II**, we set out to experimentally determine the volatilities of highly oxygenated organic molecules. We did this by utilizing a continuous stirred-tank reactor (CSTR), two cubic metres in volume and made out of teflon film. In a CSTR, reactants are continuously injected into the chamber, and air is drawn out at the same rate as the injection. This is in contrast to a batch mode reactor, which is first filled, and then air is sampled until the chamber is more or less emptied.

We injected α-pinene and ozone into the chamber to form HOMs, measured with a CI-APi-TOF. In some experiments we injected NO_2 as well, which was photolyzed to NO. This was in order to form organic nitrates upon the termination of RO_2 radicals with NO. In our experiments, we used such flow rates that the turnover time of the air in the chamber was around 40 minutes. This means that, with a constant injection of reactants and steady conditions, the reactants will reach a steady level in around two hours. After this, given that the injection is maintained at a constant level, their concentrations will stay stable. Thus, the oxidation rate of α-pinene in the chamber will be constant. Any oxidation products will therefore have a constant production rate. If they are of low enough volatility, such as ELVOCs, they will condense on chamber walls and any available aerosol surfaces essentially irreversibly. Typically, the wall loss will be their main sink: under these conditions, their concentration, to a first approximation, is determined by their chemical production rate in the gas phase and the wall loss rate. The wall loss corresponds to a lifetime in the gas phase of a couple of minutes, meaning that it typically takes this long for a low volatility compound to hit a wall after it is produced. Thus, even low-volatile compounds can have a substantial concentration in the gas phase, provided that their source term is large enough.

In addition to the gaseous reactants, we also periodically injected inorganic aerosol particles into the chamber. These seed particles increase the surface area available for condensation in the chamber, described by the condensation sink (CS). With a high enough concentration of seed particles, the dominant loss mechanism of low volatility vapours is shifted from the wall loss to the condensation on the seed particles. As this increases the total loss rate of low volatility vapours, and decreases their lifetime in the gas phase to around a minute, their concentration in the gas phase drops. We can measure this drop with the CI-APi-TOF. Compounds of comparatively higher volatility, such as SVOCs, will not condense irreversibly. Instead, they will condense,

but also evaporate back to the gas phase: therefore, the effective sink term caused by the CS is smaller for them. As a result, their concentration in the gas phase will not decrease as much, if at all, upon seed injection. To quantify the relation between the volatility of a compound, and its behaviour upon seed injection, we simulated the experiments with the ADCHAM model (Roldin et al., 2019; Roldin et al., 2014). We found that LVOCs and ELVOCs behave as if non-volatile, while SVOCs range from almost non-volatile behaviour to being almost unaffected by the seed addition.

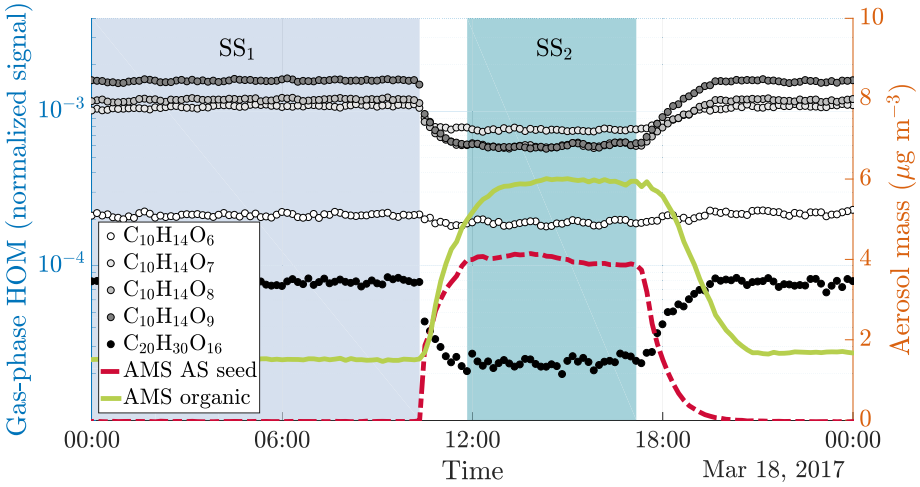

Figure 7: Time series of both gas and particle phase species during a typical seed injection. The experiment starts with only gaseous precursors being injected into the chamber: these form HOMs (dotted lines), which form SOA (solid green line). Both the HOMs and the SOA are in a steady state (SS_1, blue shading). At around 10 o'clock, we start injecting ammonium sulfate seed aerosol (red dashed line). This results in an increased condensation sink, which causes a decrease in the gas phase HOM signals, and an increase in the SOA mass. After a transition period, a new steady state is reached (SS_2, turquoise shading). Upon stopping the seed injection, the SOA levels decrease again, and the HOM signals increase. Figure from **Paper II**.

HOM monomers generally showed a progression towards irreversible condensation with increasing oxygen numbers (Fig. 7). To investigate this in more detail, we calculated the fractions remaining (FR) for each compound as a ratio of their concentrations during steady states 2 and 1, as seen in Fig. 7. The picture is broadly similar as when looking at individual compounds: at low masses, many compounds are unaffected by the seed

addition, but with an increasing mass, they start behaving as effectively non-volatile (Fig. 8). This is consistent with a decrease in volatility with increasing mass. When the compounds are of a low enough volatility, further reductions in their volatility do not change their behaviour: there is already virtually zero evaporation back to gas phase. Therefore, HOM dimers, presumably orders of magnitude less volatile than even the most oxygenated monomers, behave essentially the same (Fig. 8).

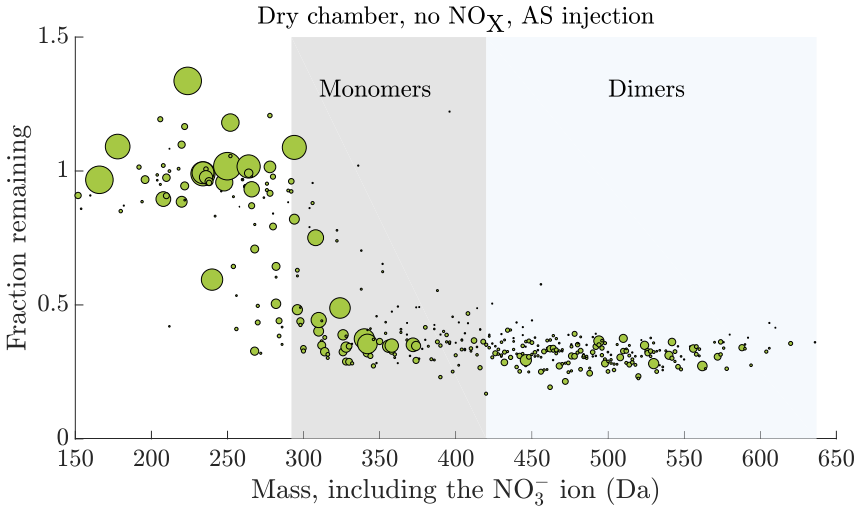

Figure 8: Fraction remaining after seed injection vs. the molecular mass of the detected cluster. The area of the circles has been linearly scaled to the magnitude of the signal of each compound before the seed injection. The data are the average of experiments with ammonium sulfate injection, dry chamber, and no NO_X in the chamber. Characteristic HOM monomer and dimer regions have been highlighted with grey and blue shading, respectively. Figure from **Paper II**.

One potential pitfall of our method for probing the volatilities of HOMs is the possibility that particle phase reactions affect the uptake of HOMs. This would cause reduced evaporation from the particles, and lead to the compounds seeming artificially low in volatility. We tried this hypothesis by using two different types of seed aerosol particles: ammonium sulfate (AS), and the more acidic ammonium bisulfate (ABS). Particle phase reactions are expected to be faster in the ABS as compared to the AS (e.g. Gao et al., 2004; Iinuma et al., 2004; Riva, Heikkinen, et al., 2019). We did not observe any significant differences in the behaviour of HOMs between the AS and ABS experiments, indicating that particle phase reactions play a minor role in determining

the behaviour of HOMs. However, we did observe a difference between the experiments conducted in a dry and humid conditions, with enhanced uptake in the humid case. This is possibly related to solubility-driven uptake of HOMs to the particles. To focus on the volatility of HOMs, we chose the dry experiments with AS for further analysis.

We also conducted experiments with NO in the chamber, to enable the formation of organic nitrates. As compared to non-nitrate HOMs, organic nitrates of similar volatility had a higher mass. This is consistent with the high mass of the nitrate group: it is expected to lower the volatility of a compound around as much as a hydroxide or a hydroperoxide group, both having a much lower mass (Kroll & Seinfeld, 2008; Pankow & Asher, 2008). To investigate how the volatility of a compound depends on its chemical composition in more detail, we constructed a logistic model, explaining the fraction remaining in terms of the number of carbon, hydrogen, oxygen and nitrogen atoms in a compound. The fraction remaining could be relatively well predicted using the molecular composition. Combining this relation with the relation between volatility and fraction remaining from the ADCHAM model, we could construct a parametrization for the volatility of a compound based on its composition:

$$\log_{10}(\mathrm{C}^*[\mu\mathrm{g\ m}^{-3}]) = 0.18 \times n_\mathrm{C} - 0.14 \times n_\mathrm{H} - 0.38 \times n_\mathrm{O} + 0.80 \times n_\mathrm{N} + 3.1. \quad (11)$$

Counter-intuitively, the coefficient for carbon in the equation is positive, indicating that adding carbon to a molecule would increase its volatility. However, this apparent fallacy is explained by the fact that typically, carbon addition is in the form of CH_2: combining effects of one carbon and two hydrogens in the equation leads to an overall negative coefficient. The same holds for nitrogen, added in the form of a nitrate group, NO_3. For HOMs of a given carbon number, there are multiple possible hydrogen numbers. For ozonolysis products, the hydrogen number is primarily determined in the RO_2 termination step, with lower numbers corresponding to the formation of carbonyls, and higher to hydroxy groups (Sect. 2.1, e.g. Eq. (4)). Hydroxy groups are expected to lower the volatility of a compound considerably more than than carbonyls (Pankow & Asher, 2008). Thus, the negative coefficient of the hydrogen number in the equation probably reflects the effect of different functional groups.

As noted before, as the volatility of a compound decreases below a certain threshold, its behaviour doesn't change any more. This means that the data the model is fitted

to loses any detailed information on volatility in this range. Therefore, the model is probably best fit for compounds in the SVOC-LVOC range, including HOM monomers. Therefore the estimates for HOM dimers, for example, may be very inaccurate.

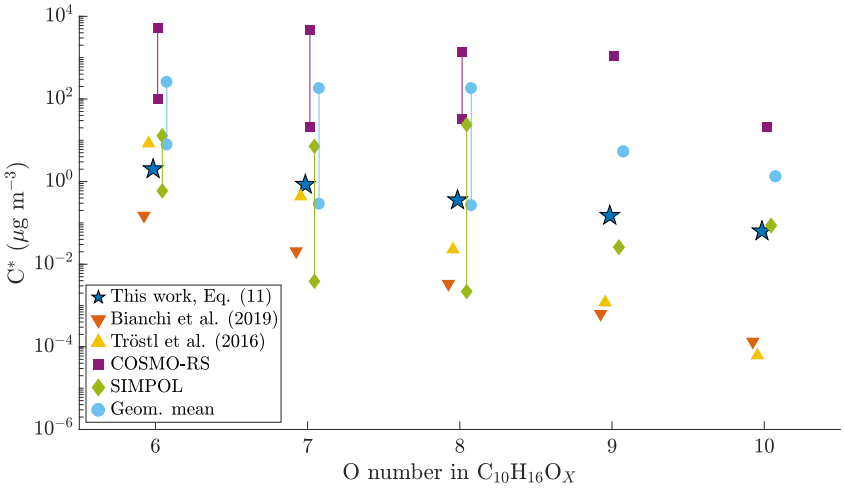

Figure 9: Comparison of different volatility estimates and parametrizations. SIMPOL and COSMO-RS are from Kurtén et al. (2016). Geom. mean is the geometric mean of the two, as recommended by Kurtén et al. (2016). Figure from **Paper II**.

Compared to other volatility parametrizations for HOMs, such as the one by Bianchi et al. (2019), the volatility of a HOMs according to Eq. (11) is much less sensitive to the addition of oxygen. In addition, while the Bianchi et al. (2019) parametrization is more sensitive to the addition of carbon than it is to oxygen, Eq. (11) has this the other way around. Thus, Eq. (11) is quite different from existing volatility parametrizations. This probably results from the fact that Eq. (11) is mainly fit for HOM monomers: their volatility dependence may be rather different from that of other types of compounds. As an example of this, Kurtén et al. (2016) suggested that HOM monomers are much more volatile than previously thought, with the volatility being much less sensitive to oxygen addition. Our results suggest a similar sensitivity to oxygen addition, but lower volatilities (Fig. 9). Thus, our results fit in between the parametrizations by Bianchi et al. (2019) and Tröstl et al. (2016), and the results by Kurtén et al. (2016) (Fig. 9). As such, they agree with previous results that while HOM monomers contribute to the growth of particles, it is the dimers that can be important for the very first steps of particle formation (Ehn et al., 2014; Kirkby et al., 2016; Tröstl et al., 2016).

6 Insights into atmospheric oxidation using a novel factorization technique

Atmospheric mass spectrometry can provide a wealth of data, often very time consuming to analyze (Sect. 3.2). Extracting the maximum amount of information from the measured spectra typically involves high resolution (HR) peak fitting. As an example, each circle in Fig. 8, taken from **Paper II**, represents a HR fitted compound: identification of the compounds from the spectra, and the validation of their fits is no small task. Further, in **Paper II** we used a mass spectrometer with a relatively high mass resolution: often, the resolution available is poorer, making HR fitting even more challenging. One approach is to sum all the signal at a given integer mass together: this unit mass resolution (UMR) analysis is less time consuming and subjective than HR fitting, but also loses much of the information contained in the spectra. Both UMR and HR data have been successfully used with different dimensionality reduction techniques, such as positive matrix factorization (PMF), to identify signals originating from similar sources (Massoli et al., 2018; Paatero & Tapper, 1994; Yan et al., 2016).

To simplify the analysis of HR data, and reduce its subjectivity, we developed a new data analysis method, mass spectral binning combined with positive matrix factorization (binPMF, **Paper III**). binPMF requires no prior information of the measured spectra. We subsequently used the method, in combination with analysis of sub-ranges of mass spectra, to identify new and known HOM formation pathways in **Paper IV**.

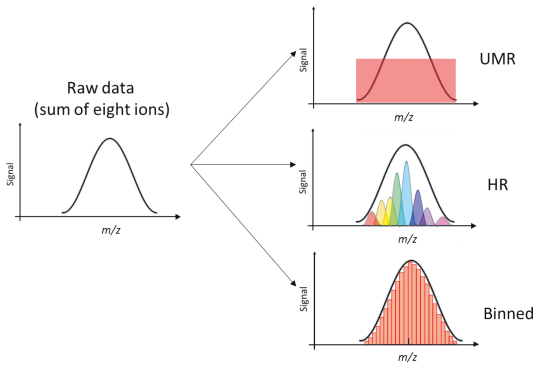

Figure 10: The concept of binning the mass spectra, compared to UMR and HR approaches. Figure from **Paper III**.

In binPMF, the mass spectrum at each time point is divided into a large number of narrow bins (Fig. 10). By utilizing a narrow enough bin width, most of the high resolution information is preserved, without the need for the error prone HR fitting. The data thus consists of the signal intensities at each bin, for each time point. Then, we use PMF to find bins that show a similar temporal variability, to combine the bins into factors, as described in Sect. 3.2.2 and visualized in Fig. 4. As in PMF in general, the data is split into a predefined number of factors, each characterized by its mass spectral profile and time behaviour. If the factorization is successful, the spectrum of any single time point can be well described as a sum of the factor profiles, weighted by their contributions at that time point.

Optimally, the individual ions at a given unit mass will be separated into their own factors, as in HR fitting. In addition, similarly behaving ions across unit masses will be grouped together, as in traditional UMR or HR PMF. Thus, binPMF can handle both the peak separation, and the assignment of peaks into factors. This can be done with no prior information of the compounds present in the mass spectrum, or of the number of compounds present at a given unit mass. As a result, binPMF greatly simplifies the data analysis workflow, at the same time making the process less subjective.

To test whether binPMF works in practice, we generated a synthetic data set containing two overlapping mass spectral peaks, including noise. The peaks were given time behaviours independent of each other, and a random mass calibration error was introduced to each time point individually. We conducted multiple such experiments, varying the peak separation between experiments.

We found that binPMF could resolve the time behaviours of the two peaks at least as well as HR fitting. With decreasing peak separation, binPMF got progressively better in relation to HR fitting. The peak positions could be accurately retrieved from the factor profiles. The performance of binPMF further improved when we introduced two additional, well separated peaks, each one correlated with one of the first two. This is representative of real atmospheric measurements: most often multiple compounds share a similar source, leading to similar time behaviours. In contrast, the performance of traditional HR fitting was unaffected by the introduction of additional peaks, as the fitting procedure does not utilize information found on other unit masses.

With the success of the synthetic dataset, we proceeded to utilize binPMF on ambient CI-APi-TOF measurements in Hyytiälä. To assess its performance, we compared the

results to those from UMR PMF. Yan et al. (2016) have previously identified multiple HOM formation pathways using UMR PMF on CI-APi-TOF data measured in Hyytiälä, giving us a point of comparison.

The results from binPMF and UMR PMF generally agreed well with each other (Fig. 11). Some factors, such as numbers 1 and 2, representing night-time and daytime HOM formation, respectively, had virtually identical mass spectral profiles and time series between the binPMF and UMR PMF solutions. For these factors, the performance of binPMF matched that of the UMR PMF, with the added benefit of the additional HR information preserved in the binPMF factor profiles. Some of the factors differed more between approaches, and certain features were only found in the binPMF results. This was the case for factor 6: only binPMF separated a factor with a clear sawtooth pattern time behaviour. This was attributed to the automatic background measurement of the instrument, inadvertently introducing fluorinated compounds to the sampling line. While scientifically uninteresting, the separation of this factor proves that binPMF can extract information from the spectrum that UMR PMF is unable to find.

For the CI-APi-TOF data measured in Hyytiälä, a large fraction of the unit masses often contain only one dominant ion. However, some masses contain multiple closely lying ions. It is for these masses that binPMF can be considered the most useful. As an example, there are at least three distinct ions that appear to be separated by binPMF on the mass 339 Th (Fig. 12). At the point of time shown in Fig. 12, only one peak is evident: $C_{10}H_{15}O_8NNO_3^-$, also identified in previous studies (Ehn et al., 2014; Kulmala et al., 2013; Yan et al., 2016). If this spectrum would have been used for the identification of peaks for HR fits, the two other peaks would have been missed. Out of the two others, that of $C_{10}H_{13}O_9NO_3^-$ also lies too close to $C_{10}H_{15}O_8NNO_3^-$ to be reliably separated by HR fitting. The third peak, contaminant $C_7HF_{10}O_4^-$, is well separated from the other two, but for the majority of the time only contributes a small fraction to the signal at this mass, and would thus be easy to miss in HR peak identification.

If the assignment of a fraction of the signal at 339 Th to the peak at $C_{10}H_{13}O_9NO_3^-$ is sound, it presents an interesting observation in terms of HOM formation. Due to the tentative assignment of the composition, this was not developed further in **Paper III**. However, I will briefly discuss the compound here, as it presents a potential connection to **Paper I**. The dominant peak at 339 Th, $C_{10}H_{15}O_8NNO_3^-$, is attributed to the organic nitrate $C_{10}H_{15}O_8N$, clustered with the nitrate ion from the CI inlet. This compound

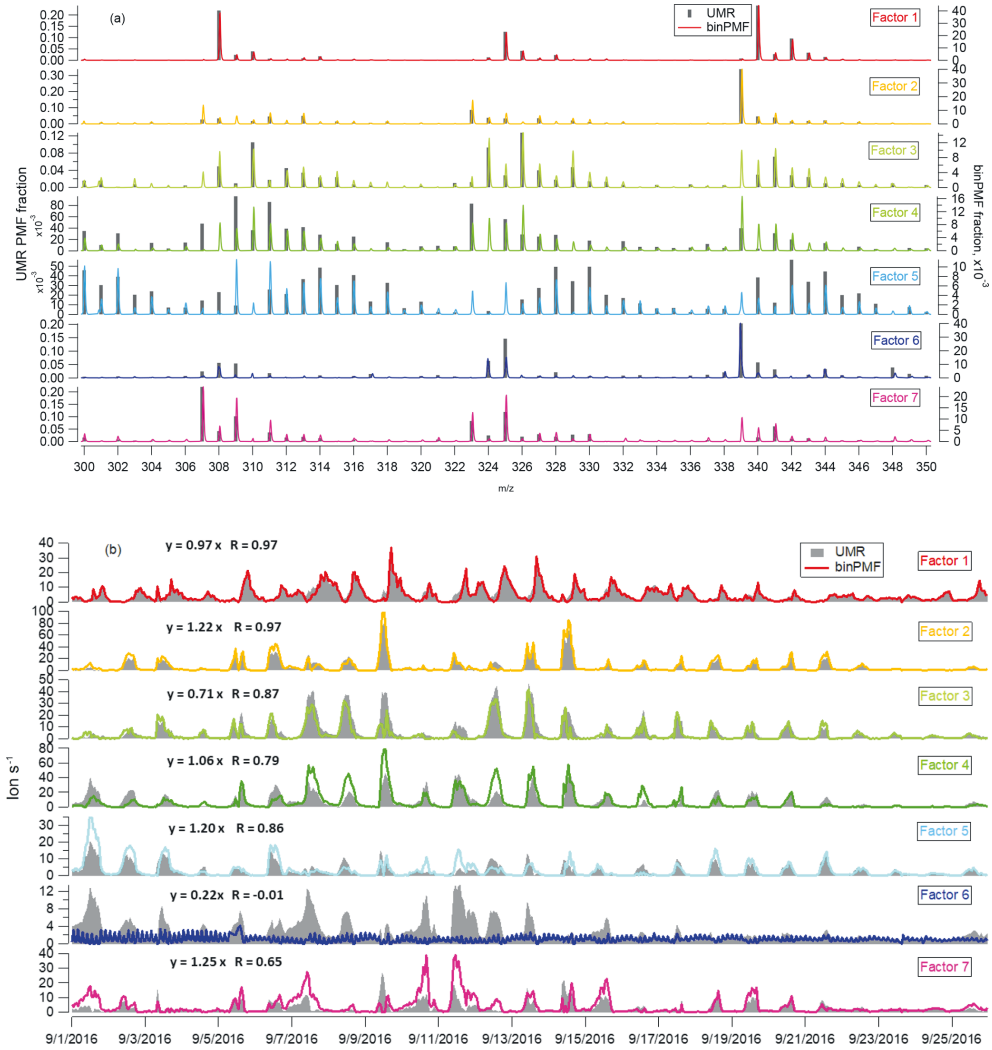

Figure 11: Comparison of binPMF and UMR PMF for factor mass spectral profiles **(a)** and time series **(b)**. The equations in each panel describe how signals from binPMF (y) compare with the UMR PMF solution (x). R is the correlation coefficient between the time series. Figure from **Paper III**.

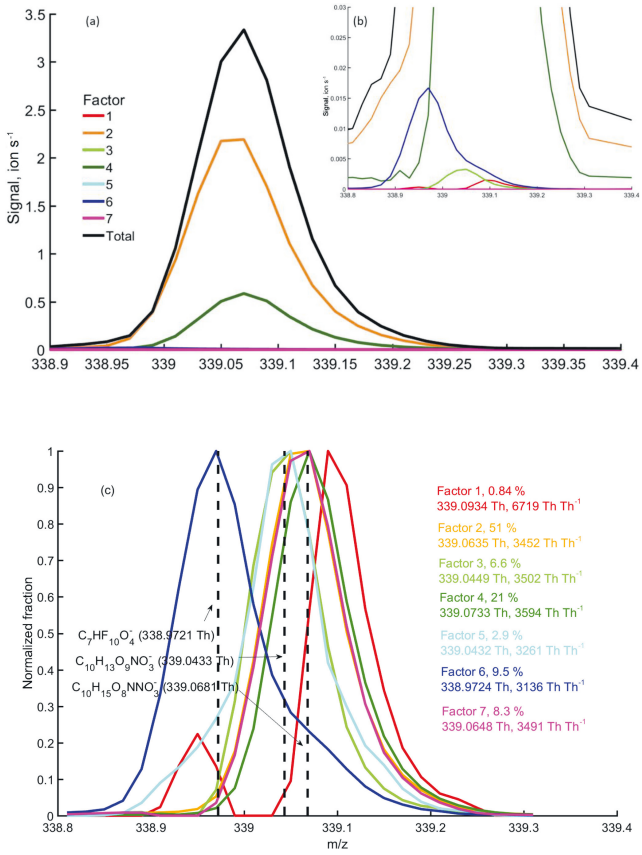

Figure 12: Factor profiles from the binPMF on mass 339 Th at 12:00 on September 9$^{\text{th}}$. Shown are both absolute contributions in **(a)** and **(b)**, and profiles normalized to a height of one in **(c)**. The location (Th) of a peak fitted to the profile and the apparent resolution (ThTh^{1-}) for each factor are given in panel **(c)**. Factor 1 has a very small contribution at this mass, and its profile is most likely erroneous at this mass. Figure from **Paper III**.

can be formed in the termination of RO_2 radicals generated in monoterpene oxidation by NO, and is thus commonly observed in Hyytiälä during the daytime (e.g. Yan et al., 2016). The other potential HOM compound at 339 Th, $C_{10}H_{13}O_9NO_3^-$, appears to be the RO_2 radical $C_{10}H_{13}O_9$, clustered with the nitrate ion. The radical has ten carbon atoms, indicating monoterpene origin. However, the radical has only thirteen hydrogen atoms: this number, three fewer than in the assumed precursor VOC, is not consistent with any of the main oxidation pathways discussed in Sect. 2.1. The minimum expected number of hydrogen atoms for a C_{10} RO_2 radical would be 15, from ozonolysis. However, many closed shell products of the ozonolysis of monoterpenes have the formula $C_{10}H_{14}O_X$. Such an oxidation product would only need to lose one hydrogen to get to the number thirteen. This would be consistent with its reaction with the OH radical by hydrogen abstraction. Such a link was proposed in **Paper I**: monoterpenes would be oxidized by ozone during the night, leading to a buildup of relatively volatile oxidation products. These could then be oxidized by OH after the sunrise, with the oxidation products contributing to the growth of particles during NPF events. The early morning peak (Fig. 8 of the **Paper III**) of factor 3, the largest contributor to the signal attributed to $C_{10}H_{13}O_9NO_3^-$, would be consistent with this pathway. However, these considerations are highly speculative, and should be considered as such until further proof emerges.

Commonly, the number of variables supplied to PMF is maximized by using a wide range of masses as input. As an example, Yan et al. (2016) selected the mass range 201 – 650 Th as input. This allows the algorithm to look for connections among a large number of compounds, but also presents some problems. As an example, the HOM dimer signals are often much lower than those of HOM monomers, even though the dimers are of key importance for particle formation. As a result, the PMF solution may be dominated by the behaviour of the monomers, missing small, but important, signal variations in the dimer range. In addition, the formation pathways of monomers and dimers can be fundamentally different. While both can be formed in RO_2 + RO_2 reactions, the identity of both participating RO_2 is directly reflected in the composition of the resulting dimer, but not the monomers. Thus, while they share similar sources, HOM monomers and dimers are not expected to be fully correlated. To illustrate the point further, RO_2 generated from the oxidation of monoterpenes can also dimerize with RO_2 from isoprene oxidation, producing C_{15} dimers (Berndt, Mentler, et al., 2018). No directly corresponding monomers exist: each can be directly attributed to either monoterpene or isoprene oxidation, not both.

To address these limitations, we split the mass spectrum into subranges, which we analyzed separately in **Paper IV**. As a comparison, we also analyzed the combined range as one. We found that we could gain new insight into the dimer formation processes, which was lost when analysing a wider range at the same time. When analysing the combined mass range, the factor separation was mainly driven by the higher signals at lower masses. The dimer signals at higher masses were allocated to the time behaviours determined by the low masses. In contrast, when analysing the dimers separately, the factorization was purely driven by their variation, allowing us to extract more information on dimer formation.

When analysing the dimer range separately, we were able to find two night-time dimer factors, and one daytime dimer factor (Fig. 13). In contrast, from the combined range, only one night-time factor was found. This showed a clear correlation with the first night-time factor of the dimer range, and could be attributed to monoterpene ozonolysis as described by Ehn et al. (2014): a similar factor was also found by Yan et al. (2016).

The second night-time dimer factor did not have counterparts in any of the other analyzed ranges, including the combined range. Already this highlights the need to analyze the ranges separately. This factor, characterized by dimers containing one nitrogen atom, was attributed to dimer formation from one RO_2 originating in ozonolysis, and one in NO_3 radical initiated oxidation. Yan et al. (2016) found a similar factor, attributed to NO_3 oxidation, but with a considerable contribution from HOM monomers as well. There are a few possibilities explaining this difference. One possibility is that in Hyytiälä, NO_3 initiated oxidation does not lead to the formation of HOM monomers. However, it does produce RO_2 radicals, which can react with the more oxygenated RO_2 from ozonolysis. Thus, the dimerization would enable them to contribute to the formation of condensible vapours, when NO_3 oxidation alone would not produce them. The reason for the monomers being attributed to the NO_3 factor by Yan et al. (2016) could be an artefact from analysing the whole mass range simultaneously. It is also possible that the factor assignment by Yan et al. (2016) is correct. One explanation could be the difference in the time of the year: the measurements used by Yan et al. (2016) were conducted in the spring, while ours were in the autumn. The monoterpene mixture in Hyytiälä varies with season (Hakola et al., 2012): thus, it is possible that the monoterpenes present in Hyytiälä during the springtime are more efficient at forming HOMs upon NO_3 oxidation than those in the autumn. Finally, during the measurements in **Paper IV**, there was frequent fog formation and depletion of ozone

close to the ground (Zha et al., 2018). This may also have impacted our results.

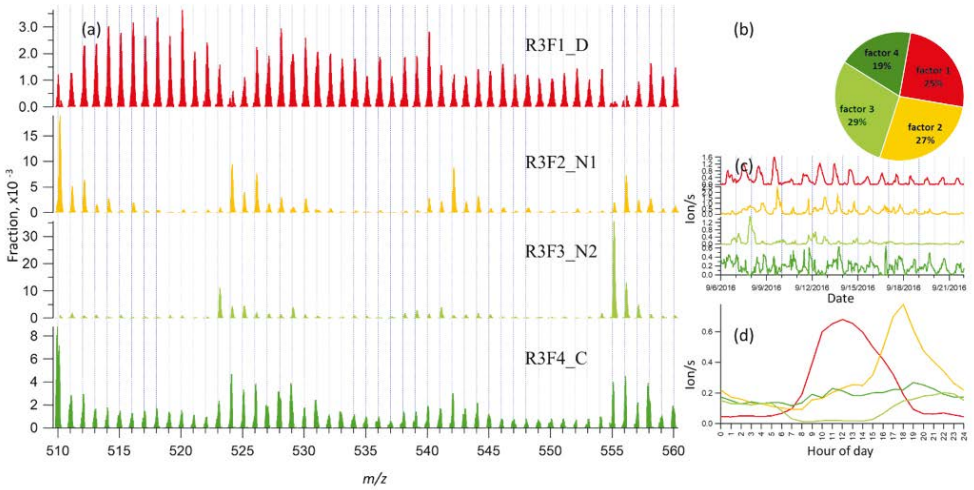

Figure 13: A four factor solution for binPMF on a mass range containing HOM dimers. **(a)**: the factor profiles, **(b)**: the contributions of each factor to the total signal, **(c)**: factor time series, and **(d)**: factor diurnal profiles. Factor 1 has a clear daytime peak, tracking mainly solar radiation. Factors 2 and 3 have night-time maxima, and factor 4 is attributed to contamination. Figure from **Paper IV**.

The daytime dimer factor had a very prominent diurnal cycle, closely tracking solar radiation. The mass spectrum of the daytime dimer factor showed relatively even contributions from multiple masses (Fig. 13). This is in contrast to the night-time factors, and HOM dimers in general. For example, the main peaks in the first night-time dimer factor are spaced either 16 or 2 Th apart. The differences come from the dimers having either 30 or 32 hydrogen atoms (2 Th difference), and the difference in oxygen numbers (16 Th difference). The lack of a clear pattern in the daytime dimer factor implies a much wider spread of element numbers in the dimers. As with the $C_{10}H_{13}O_9$ radical discussed above, this spread could potentially be caused by repeated OH oxidation. This would again be consistent with the hypothesis from **Paper I** that multi-generation OH oxidation leads to the formation of condensible vapours. As before, this remains on a highly speculative level.

7 Review of papers and the author's contribution

Paper I presents the temporal behaviour of monoterpenes and their oxidants at a boreal forests station, and finds a potential link between monoterpene oxidation and the growth of newly formed particles. I conducted all of the data analysis in the paper, apart from the preprocessing of the data and the growth rate calculations. I was also mainly responsible for the interpretation of the results, and wrote the manuscript.

Paper II determines the volatilities of HOMs in a chamber experiment. I took part in the planning of the concept of the experiment, conducted the CI-APi-TOF measurements, preprocessed and analyzed the CI-APi-TOF data, and interpreted the results. I also coordinated the collaboration to include the ADCHAM model results that connect the observations to the volatilities of HOMs, and wrote the manuscript.

Paper III introduces a new data analysis concept, binning of mass spectra combined with positive matrix factorization (binPMF). As a shared first author, I conceived the idea of using PMF on binned data. I wrote a large fraction of the analysis codes, and took part in the interpretation of the results. I also wrote parts of the manuscript.

Paper IV applies binPMF on sub-ranges of mass spectra, as opposed to maximizing the amount of data analyzed simultaneously. We find that many formation processes of HOM dimers are only found by analysing the dimers separately from the rest of the data. I took part in the planning of the study, interpretation of the results, and writing of the manuscript.

8 Conclusions and outlook

In this thesis, I studied the formation of condensible vapours from monoterpene oxidation, using both ambient and laboratory measurements. In the beginning of my PhD work, I set out with three objectives, as outlined in the introduction of this thesis:

1. Determining how the oxidation of monoterpenes by the main oxidants varies at a boreal forest station, and its connection to the growth of newly formed particles.
2. Experimentally characterising the volatilities of HOMs, and establishing what determines them.
3. Developing a new data analysis method for mass spectrometry to gain new insights into the monoterpene oxidation pathways.

I addressed the first objective in **Paper I** using long term measurements at the boreal forest station of Hyytiälä. We established that when monoterpenes are abundant, ozone dominates their night-time oxidation, with nitrate radical contributing substantially as well. The daytime oxidation of monoterpenes was dominated by ozone or OH, depending on the time of year. We also found the daytime OH oxidation to be important for the growth of newly formed particles, with indications of multi-generation oxidation producing low volatility vapours. While inconclusive, we observed compounds consistent with this hypothesis in **Papers III** and **IV**. In **Paper IV**, we found two night-time HOM dimer formation pathways, monoterpene ozonolysis and ozonolysis together with nitrate radical oxidation, consistent with the night-time findings of **Paper I**.

The second objective was addressed in **Paper II** using laboratory experiments. I observed that, as expected, increasing oxygen numbers in the HOMs led to enhanced condensation on inorganic seed particles. The most highly oxygenated HOMs were condensing essentially irreversibly. We found that most of the HOM monomers are likely LVOCs, but a considerable fraction showed semi-volatile behaviour. In contrast, all of the HOM dimers were either LVOCs, or possibly ELVOCs. We developed a parametrization for the volatilities of HOMs based on our results. Our parametrization predicts a higher volatility, and a lower sensitivity to oxygen addition, than many existing ones. Our findings are similar to recent computational results: however, our parametrization predicts lower absolute values for the volatilities. All in all, our results

support the recent views that HOM monomers are efficient at forming SOA, but only the HOM dimers can take part in the very earliest stages of new particle formation.

To address the third objective, we developed a new data analysis method, binPMF, in **Paper III**. We further applied the method in **Paper IV** on HOM monomers and dimers separately, as opposed to simultaneously, as is commonly done. In binPMF, we divide the mass spectrum into narrow bins, and use the time series of each bin as input for PMF. PMF then splits the data into distinct factors. binPMF has clear advantages as compared to both traditional peak fitting methods and to UMR PMF. Unlike the former, binPMF requires minimal prior information of the measured spectra. However, it still conserves high resolution information, lost in the latter. This additional information allows binPMF to identify more factors than UMR PMF, and enables better characterization of the factors found by both. Using binPMF separately on HOM monomers and dimers in **Paper IV**, we were able to separate three different types of HOM dimers: two night-time, and one daytime type. These were not discovered by analysing monomers and dimers simultaneously.

The link between the preceding night and the growth of new particles during the next day, proposed in **Paper I**, remains uncertain. Long term measurements of HOMs in Hyytiälä could possibly clarify the link. Our parametrization for the volatilities of HOMs formed in the ozonolysis of α-pinene, developed in **Paper II**, differs considerably from earlier parametrizations, predicting generally higher volatilities. This affects the role HOMs play in the early growth of newly formed particles. Comparing the parametrizations in terms of how they match with observations of early particle growth would be an interesting topic for further study. binPMF, introduced in **Paper III** and applied further in **Paper IV**, should prove useful in the analysis of a wide variety of mass spectrometric data. Particle phase measurements, as well as gas phase measurements using ionization schemes other than nitrate, would be logical next applications. Some refinements, such as improved selection of the bin width, could potentially further enhance the performance of binPMF.

Taken together, the results of this thesis are in line with prior findings that monoterpene oxidation is a key contributor to the formation of SOA and the growth of newly formed particles. These results provide insight into the ways in which the particle formation occurs, and shed light on some of the current key questions regarding monoterpene oxidation, such as the volatilities of HOMs. This work should offer ample ground for further research into VOC oxidation and aerosol formation in the future.

References

Allan, B., McFiggans, G., Plane, J., Coe, H., & McFadyen, G. (2000). The nitrate radical in the remote marine boundary layer. *Journal of Geophysical Research-Atmospheres*, *105*(D19), 24191–24204. doi:10.1029/2000JD900314

Andreae, M. O. (2007). Aerosols before pollution. *Science*, *315*(5808), 50–51. doi:10.1126/science.1136529

Atkinson, R., & Arey, J. (2003). Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review. *Atmospheric Environment*, *37*(2), S197–S219. doi:10.1016/S1352-2310(03)00391-1

Bäck, J., Aalto, J., Henriksson, M., Hakola, H., He, Q., & Boy, M. (2012). Chemodiversity of a Scots pine stand and implications for terpene air concentrations. *Biogeosciences*, *9*(2), 689–702. doi:10.5194/bg-9-689-2012

Barsanti, K. C., Smith, J. N., & Pankow, J. F. (2011). Application of the np + mP modeling approach for simulating secondary organic particulate matter formation from α-pinene oxidation. *Atmospheric Environment*, *45*(37), 6812–6819. doi:10.1016/j.atmosenv.2011.01.038

Berndt, T., Mentler, B., Scholz, W., Fischer, L., Herrmann, H., Kulmala, M., & Hansel, A. (2018). Accretion product formation from ozonolysis and OH radical reaction of α-pinene: Mechanistic insight and the influence of isoprene and ethylene. *Environmental Science & Technology*, *52*(19), 11069–11077. doi:10.1021/acs.est.8b02210

Berndt, T., Richters, S., Jokinen, T., Hyttinen, N., Kurtén, T., Otkjær, R. V., ... Ehn, M. (2016). Hydroxyl radical-induced formation of highly oxidized organic compounds. *Nature Communications*, *7*(1). doi:10.1038/ncomms13677

Berndt, T., Richters, S., Kaethner, R., Voigtländer, J., Stratmann, F., Sipilä, M., ... Herrmann, H. (2015). Gas-phase ozonolysis of cycloalkenes: Formation of highly oxidized RO_2 radicals and their reactions with NO, NO_2, SO_2, and other RO_2 radicals. *The Journal of Physical Chemistry A*, *119*(41), 10336–10348. doi:10.1021/acs.jpca.5b07295

Berndt, T., Scholz, W., Mentler, B., Fischer, L., Herrmann, H., Kulmala, M., & Hansel, A. (2018). Accretion product formation from self- and cross-reactions of RO_2 radicals in the atmosphere. *Angewandte Chemie Int Ed*, *57*(14). doi:10.1002/anie.201710989

Bianchi, F., Kurtén, T., Riva, M., Mohr, C., Rissanen, M. P., Roldin, P., ... Ehn, M. (2019). Highly oxygenated organic molecules (HOM) from gas-phase autoxidation

involving peroxy radicals: A key contributor to atmospheric aerosol. *Chemical Reviews*, *119*(6). doi:10.1021/acs.chemrev.8b00395

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... Zhang, X. (2013). Clouds and aerosols. In T. F. Stocker, D. Qin, G.-K. Plattner, M. M. Tignor, S. K. Allen, J. Boschung, ... P. M. Midgley (Eds.). doi:10.1017/CBO9781107415324

Breitenlechner, M., Fischer, L., Hainer, M., Heinritzi, M., Curtius, J., & Hansel, A. (2017). PTR3: An instrument for studying the lifecycle of reactive organic carbon in the atmosphere. *Analytical Chemistry*, *89*(11), 5824–5831. doi:10.1021/acs.analchem.6b05110

Carslaw, K., Lee, L., Reddington, C., Pringle, K., Rap, A., Forster, P., ... Pierce, J. (2013). Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, *503*(7474), 67–71. doi:10.1038/nature12674

Chen, D., Zhou, P., Nieminen, T., Roldin, P., Qi, X., Clusius, P., ... Boy, M. (2020). The trend of the oxidants in boreal forest over 2007–2018: Comprehensive modelling study with long-term measurements at SMEAR II, Finland. *Atmospheric Chemistry and Physics Discussions*, *2020*, 1–41. doi:10.5194/acp-2020-128

Cox, R., & Cole, J. (1985). Chemical aspects of the autoignition of hydrocarbon-air mixtures. *Combustion and Flame*, *60*(2), 109–123. doi:10.1016/0010-2180(85)90001-x

Crounse, J. D., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G., & Wennberg, P. O. (2013). Autoxidation of organic compounds in the atmosphere. *The Journal of Physical Chemistry Letters*, *4*(20), 3513–3520. doi:10.1021/jz4019207

Cubison, M., & Jimenez, J. (2015). Statistical precision of the intensities retrieved from constrained fitting of overlapping peaks in high-resolution mass spectra. *Atmospheric Measurement Techniques*, *8*(6), 2333–2345. doi:10.5194/amt-8-2333-2015

Donahue, N. M., Epstein, S. A., Pandis, S. N., & Robinson, A. L. (2011). A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics. *Atmospheric Chemistry and Physics*, *11*(7), 3303–3318. doi:10.5194/acp-11-3303-2011

Donahue, N. M., Kroll, J. H., Pandis, S. N., & Robinson, A. L. (2012). A two-dimensional volatility basis set – part 2: Diagnostics of organic-aerosol evolution. *Atmospheric Chemistry and Physics*, *12*(2), 615–634. doi:10.5194/acp-12-615-2012

Donahue, N. M., Trump, E. R., Pierce, J. R., & Riipinen, I. (2011). Theoretical constraints on pure vapor-pressure driven condensation of organics to ultrafine particles. *Geophysical Research Letters*, *38*(16), n/a–n/a. doi:10.1029/2011gl048115

Ehn, M., Junninen, H., Petäjä, T., Kurtén, T., Kerminen, V.-M., Schobesberger, S., ... Worsnop, D. R. (2010). Composition and temporal behavior of ambient ions in the boreal forest. *Atmospheric Chemistry and Physics*, *10*(17), 8513–8530. doi:10.5194/acp-10-8513-2010

Ehn, M., Kleist, E., Junninen, H., Petäjä, T., Lönn, G., Schobesberger, S., ... Mentel, T. F. (2012). Gas phase formation of extremely oxidized pinene reaction products in chamber and ambient air. *Atmospheric Chemistry and Physics*, *12*(11), 5113–5127. doi:10.5194/acp-12-5113-2012

Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., ... Mentel, T. F. (2014). A large source of low-volatility secondary organic aerosol. *Nature*, *506*(7489), 476–479. doi:10.1038/nature13032

Eisele, F. L., & Tanner, D. J. (1993). Measurement of the gas phase concentration of H_2SO_4 and methane sulfonic acid and estimates of H_2SO_4 production and loss in the atmosphere. *Journal of Geophysical Research: Atmospheres*, *98*(D5), 9001–9010. doi:10.1029/93jd00031

Etheridge, D. M., Steele, L. P., Langenfelds, R. L., Francey, R. J., Barnola, J.-M., & Morgan, V. I. (1996). Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn. *Journal of Geophysical Research: Atmospheres*, *101*(D2), 4115–4128. doi:10.1029/95jd03410

Gakidou, E., Afshin, A., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., ... Murray, C. J. L. (2017). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: A systematic analysis for the global burden of disease study 2016. *The Lancet*, *390*(10100), 1345–1422. doi:10.1016/s0140-6736(17)32366-8

Gao, S., Ng, N. L., Keywood, M., Varutbangkul, V., Bahreini, R., Nenes, A., ... Seinfeld, J. H. (2004). Particle phase acidity and oligomer formation in secondary organic aerosol. *Environmental Science & Technology*, *38*(24), 6582–6589. doi:10.1021/es049125k

Garmash, O., Rissanen, M. P., Pullinen, I., Schmitt, S., Kausiala, O., Tillmann, R., ... Ehn, M. (2020). Multi-generation OH oxidation as a source for highly oxygenated

organic molecules from aromatics. *Atmospheric Chemistry and Physics*, *20*(1), 515–537. doi:10.5194/acp-20-515-2020

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, *5*(6), 1471–1492. doi:10.5194/gmd-5-1471-2012

Hakola, H., Hellén, H., Hemmilä, M., Rinne, J., & Kulmala, M. (2012). In situ measurements of volatile organic compounds in a boreal forest. *Atmospheric Chemistry and Physics*, *12*(23), 11665–11678. doi:10.5194/acp-12-11665-2012

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., ... Wildt, J. (2009). The formation, properties and impact of secondary organic aerosol: Current and emerging issues. *Atmospheric Chemistry and Physics*, *9*(14), 5155–5236. doi:10.5194/acp-9-5155-2009

Hari, P., & Kulmala, M. (2005). Station for measuring ecosystem-atmosphere relations (SMEAR II). *Boreal Environment Research*, *10*(5), 315–322.

Hirsikko, A., Laakso, L., Hõrrak, U., Aalto, P., Kerminen, V.-M., & Kulmala, M. (2005). Annual and size dependent variation of growth rates and ion concentrations in boreal forest. *Boreal environment research*, *10*(5), 357–369.

Hyttinen, N., Knap, H. C., Rissanen, M. P., Jørgensen, S., Kjaergaard, H. G., & Kurtén, T. (2016). Unimolecular HO_2 loss from peroxy radicals formed in autoxidation is unlikely under atmospheric conditions. *The Journal of Physical Chemistry A*, *120*(20), 3588–3595. doi:10.1021/acs.jpca.6b02281

Hyttinen, N., Kupiainen-Määttä, O., Rissanen, M. P., Muuronen, M., Ehn, M., & Kurtén, T. (2015). Modeling the charging of highly oxidized cyclohexene ozonolysis products using nitrate-based chemical ionization. *The Journal of Physical Chemistry A*, *119*(24), 6339–6345. doi:10.1021/acs.jpca.5b01818

Iinuma, Y., Böge, O., Gnauk, T., & Herrmann, H. (2004). Aerosol-chamber study of the α-pinene/O_3 reaction: Influence of particle acidity on aerosol yields and products. *Atmospheric Environment*, *38*(5), 761–773. doi:10.1016/j.atmosenv.2003.10.015

IPCC. (2013). *Climate change 2013: The physical science basis. contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change* (T. F. Stocker, D. Qin, G.-K. Plattner, M. M. Tignor, S. K. Allen, J. Boschung, ... P. M. Midgley, Eds.). doi:10.1017/CBO9781107415324

IPCC. (2018). *Global warming of 1.5°C. an IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. Shukla, ... T. Waterfield, Eds.). In Press.

Jimenez, J., Canagaratna, M., Donahue, N., Prevot, A., Zhang, Q., Kroll, J., ... Worsnop, D. (2009). Evolution of organic aerosols in the atmosphere. *Science*, *326*(5959), 1525–1529. doi:10.1126/science.1180353

Jokinen, T., Berndt, T., Makkonen, R., Kerminen, V.-M., Junninen, H., Paasonen, P., ... Sipilä, M. (2015). Production of extremely low volatile organic compounds from biogenic emissions: Measured yields and atmospheric implications. *Proceedings of the National Academy of Sciences*, *112*(23), 7123–7128. doi:10.1073/pnas.1423977112

Jokinen, T., Kausiala, O., Garmash, O., Peräkylä, O., Junninen, H., Schobesberger, S., ... Rissanen, M. P. (2016). Production of highly oxidized organic compounds from ozonolysis of β-caryophyllene: Laboratory and field measurements. *Boreal Environment Research*, *21*(3-4), 262–273.

Jokinen, T., Sipilä, M., Junninen, H., Ehn, M., Lönn, G., Hakala, J., ... Worsnop, D. R. (2012). Atmospheric sulphuric acid and neutral cluster measurements using CI-APi-TOF. *Atmospheric Chemistry and Physics*, *12*(9), 4117–4125. doi:10.5194/acp-12-4117-2012

Jokinen, T., Sipilä, M., Richters, S., Kerminen, V.-M., Paasonen, P., Stratmann, F., ... Berndt, T. (2014). Rapid autoxidation forms highly oxidized RO_2 radicals in the atmosphere. *Angewandte Chemie International Edition*, *53*(52), 14596–14600. doi:10.1002/anie.201408566

Junninen, H., Ehn, M., Petäjä, T., Luosujärvi, L., Kotiaho, T., Kostiainen, R., ... Worsnop, D. R. (2010). A high-resolution mass spectrometer to measure atmospheric ion composition. *Atmospheric Measurement Techniques*, *3*(4), 1039–1053. doi:10.5194/amt-3-1039-2010

Kerminen, V.-M., Chen, X., Vakkari, V., Petäjä, T., Kulmala, M., & Bianchi, F. (2018). Atmospheric new particle formation and growth: Review of field observations. *Environmental Research Letters*, *13*(10), 103003. doi:10.1088/1748-9326/aadf3c

Kerminen, V.-M., Paramonov, M., Anttila, T., Riipinen, I., Fountoukis, C., Korhonen, H., ... Petäjä, T. (2012). Cloud condensation nuclei production associated with

atmospheric nucleation: a synthesis based on existing literature and new results. *Atmospheric Chemistry and Physics*, *12*(24), 12037–12059. doi:10.5194/acp-12-12037-2012

Kiendler-Scharr, A., Wildt, J., Maso, M. D., Hohaus, T., Kleist, E., Mentel, T. F., ... Wahner, A. (2009). New particle formation in forests inhibited by isoprene emissions. *Nature*, *461*(7262), 381–384. doi:10.1038/nature08292

Kirkby, J., Curtius, J., Almeida, J., Dunne, E., Duplissy, J., Ehrhart, S., ... Kulmala, M. (2011). Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation. *Nature*, *476*(7361), 429–433. doi:10.1038/nature10343

Kirkby, J., Duplissy, J., Sengupta, K., Frege, C., Gordon, H., Williamson, C., ... Curtius, J. (2016). Ion-induced nucleation of pure biogenic particles. *Nature*, *533*(7604), 521–526. Letter. doi:10.1038/nature17953

Krechmer, J., Lopez-Hilfiker, F., Koss, A., Hutterli, M., Stoermer, C., Deming, B., ... de Gouw, J. (2018). Evaluation of a new reagent-ion source and focusing ion–molecule reactor for use in proton-transfer-reaction mass spectrometry. *Analytical Chemistry*, *90*(20), 12011–12018. doi:10.1021/acs.analchem.8b02641

Kroll, J. H., & Seinfeld, J. H. (2008). Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere. *Atmospheric Environment*, *42*(16), 3593–3624. doi:10.1016/j.atmosenv.2008.01.003

Kulmala, M., Kontkanen, J., Junninen, H., Lehtipalo, K., Manninen, H. E., Nieminen, T., ... Worsnop, D. R. (2013). Direct observations of atmospheric aerosol nucleation. *Science*, *339*(6122), 943–946. doi:10.1126/science.1227385

Kulmala, M., Toivonen, A., Mäkelä, J., & Laaksonen, A. (1998). Analysis of the growth of nucleation mode particles observed in boreal forest. *Tellus Series B-Chemical And Physical Meteorology*, *50*(5), 449–462. doi:10.3402/tellusb.v50i5.16229

Kurtén, T., Tiusanen, K., Roldin, P., Rissanen, M., Luy, J., Boy, M., ... Donahue, N. (2016). α-pinene autoxidation products may not have extremely low saturation vapor pressures despite high O:C ratios. *J Phys Chem*, *120*(16), 2569–2582. doi:10.1021/acs.jpca.6b02196

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., ... van Vuuren, D. P. (2010). Historical (1850 – 2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application. *Atmospheric Chemistry and Physics*, *10*(15), 7017–7039. doi:10.5194/acp-10-7017-2010

Laothawornkitkul, J., Taylor, J., Paul, N., & Hewitt, C. (2009). Biogenic volatile organic compounds in the Earth system. *New Phytologist*, *183*(1), 27–51. doi:10.1111/j.1469-8137.2009.02859.x

Liebmann, J., Karu, E., Sobanski, N., Schuladen, J., Ehn, M., Schallhart, S., ... Crowley, J. N. (2018). Direct measurement of NO_3 radical reactivity in a boreal forest. *Atmospheric Chemistry and Physics*, *18*(5), 3799–3815. doi:10.5194/acp-18-3799-2018

Lindinger, W., Hansel, A., & Jordan, A. (1998). Proton-transfer-reaction mass spectrometry (PTR-MS): on-line monitoring of volatile organic compounds at pptv levels. *Chemical Society Reviews*, *27*(5), 347–354. doi:10.1039/a827347z

Lüthi, D., Floch, M. L., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., ... Stocker, T. F. (2008). High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature*, *453*(7193), 379–382. doi:10.1038/nature06949

Massoli, P., Stark, H., Canagaratna, M. R., Krechmer, J. E., Xu, L., Ng, N. L., ... Worsnop, D. R. (2018). Ambient measurements of highly oxidized gas-phase molecules during the southern oxidant and aerosol study (SOAS) 2013. *ACS Earth and Space Chemistry*, *2*(7), 653–672. doi:10.1021/acsearthspacechem.8b00028

Mauldin, R. L., III, Berndt, T., Sipilae, M., Paasonen, P., Petaja, T., Kim, S., ... Kulmala, M. (2012). A new atmospherically relevant oxidant of sulphur dioxide. *Nature*, *488*(7410), 193+. doi:10.1038/nature11278

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., ... Trainer, M. (2018). Volatile chemical products emerging as largest petrochemical source of urban organic emissions. *Science*, *359*(6377), 760–764. doi:10.1126/science.aaq0524

McFiggans, G., Mentel, T. F., Wildt, J., Pullinen, I., Kang, S., Kleist, E., ... Kiendler-Scharr, A. (2019). Secondary organic aerosol reduced by mixture of atmospheric vapours. *Nature*, *565*(7741), 587–593. doi:10.1038/s41586-018-0871-y

Mentel, T. F., Springer, M., Ehn, M., Kleist, E., Pullinen, I., Kurtén, T., ... Wildt, J. (2015). Formation of highly oxidized multifunctional compounds: Autoxidation of peroxy radicals formed in the ozonolysis of alkenes – deduced from structure–product relationships. *Atmospheric Chemistry and Physics*, *15*(12), 6745–6765. doi:10.5194/acp-15-6745-2015

Miliordos, E., & Xantheas, S. S. (2014). On the bonding nature of ozone (O_3) and its sulfur-substituted analogues SO_2, OS_2, and S_3: Correlation between their biradical character and molecular properties. *Journal of the American Chemical Society*, *136*(7), 2808–2817. doi:10.1021/ja410726u

Molteni, U., Bianchi, F., Klein, F., Haddad, I. E., Frege, C., Rossi, M. J., ... Baltensperger, U. (2018). Formation of highly oxygenated organic molecules from aromatic compounds. *Atmospheric Chemistry and Physics*, *18*(3), 1909–1921. doi:10.5194/acp-18-1909-2018

Paatero, P., & Tapper, U. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, *5*(2), 111–126. doi:10.1002/env.3170050203

Pankow, J. F., & Asher, W. E. (2008). SIMPOL.1: A simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds. *Atmospheric Chemistry and Physics*, *8*(10), 2773–2796. doi:10.5194/acp-8-2773-2008

Petäjä, T., Mauldin, R. L., III, Kosciuch, E., McGrath, J., Nieminen, T., Paasonen, P., ... Kulmala, M. (2009). Sulfuric acid and OH concentrations in a boreal forest site. *Atmospheric Chemistry and Physics*, *9*(19), 7435–7448.

Pierce, J. R., Riipinen, I., Kulmala, M., Ehn, M., Petäjä, T., Junninen, H., ... Donahue, N. M. (2011). Quantification of the volatility of secondary organic compounds in ultrafine particles during nucleation events. *Atmospheric Chemistry and Physics*, *11*(17), 9019–9036. doi:10.5194/acp-11-9019-2011

Riipinen, I., Pierce, J. R., Yli-Juuti, T., Nieminen, T., Häkkinen, S., Ehn, M., ... Kulmala, M. (2011). Organic condensation: a vital link connecting aerosol formation to cloud condensation nuclei (CCN) concentrations. *Atmospheric Chemistry and Physics*, *11*(8), 3865–3878. doi:10.5194/acp-11-3865-2011

Rinne, J., Bäck, J., & Hakola, H. (2009). Biogenic volatile organic compound emissions from the Eurasian taiga: current knowledge and future directions. *Boreal Environment Research*, *14*(4), 807–826.

Rinne, J., Markkanen, T., Ruuskanen, T. M., Petäjä, T., Keronen, P., Tang, M. J., ... Vesala, T. (2012). Effect of chemical degradation on fluxes of reactive compounds - a study with a stochastic Lagrangian transport model. *Atmospheric Chemistry and Physics*, *12*(11), 4843–4854. doi:10.5194/acp-12-4843-2012

Rissanen, M. P., Kurtén, T., Sipilä, M., Thornton, J. A., Kangasluoma, J., Sarnela, N., ... Ehn, M. (2014). The formation of highly oxidized multifunctional prod-

ucts in the ozonolysis of cyclohexene. *Journal of the American Chemical Society*, *136*(44), 15596–15606. PMID: 25283472. doi:10.1021/ja507146s

Riva, M., Heikkinen, L., Bell, D., Peräkylä, O., Zha, Q., Schallhart, S., ... Ehn, M. (2019). Chemical transformations in monoterpene-derived organic aerosol enhanced by inorganic composition. *npj Climate and Atmospheric Science*, *2*(1), 2. doi:10.1038/s41612-018-0058-0

Riva, M., Rantala, P., Krechmer, J. E., Peräkylä, O., Zhang, Y., Heikkinen, L., ... Ehn, M. (2019). Evaluating the performance of five different chemical ionization techniques for detecting gaseous oxygenated organic species. *Atmospheric Measurement Techniques*, *12*(4), 2403–2421. doi:10.5194/amt-12-2403-2019

Röhrer, F., & Berresheim, H. (2006). Strong correlation between levels of tropospheric hydroxyl radicals and solar ultraviolet radiation. *Nature*, *442*(7099), 184–187. doi:10.1038/nature04924

Roldin, P., Ehn, M., Kurtén, T., Olenius, T., Rissanen, M. P., Sarnela, N., ... Boy, M. (2019). The role of highly oxygenated organic molecules in the boreal aerosol-cloud-climate system. *Nature Communications*, *10*(1). doi:10.1038/s41467-019-12338-8

Roldin, P., Eriksson, A. C., Nordin, E. Z., Hermansson, E., Mogensen, D., Rusanen, A., ... Pagels, J. (2014). Modelling non-equilibrium secondary organic aerosol formation and evaporation with the aerosol dynamics, gas- and particle-phase chemistry kinetic multilayer model ADCHAM. *Atmospheric Chemistry and Physics*, *14*(15), 7953–7993. doi:10.5194/acp-14-7953-2014

Rose, C., Zha, Q., Dada, L., Yan, C., Lehtipalo, K., Junninen, H., ... Kulmala, M. (2018). Observations of biogenic ion-induced cluster formation in the atmosphere. *Science Advances*, *4*(4). doi:10.1126/sciadv.aar5218

Rostami, A. A. (2009). Computational modeling of aerosol deposition in respiratory tract: A review. *Inhalation Toxicology*, *21*(4), 262–290. doi:10.1080/08958370802448987

Schobesberger, S., Junninen, H., Bianchi, F., Lönn, G., Ehn, M., Lehtipalo, K., ... Worsnop, D. R. (2013). Molecular understanding of atmospheric particle formation from sulfuric acid and large oxidized organic molecules. *Proceedings of the National Academy of Sciences*, *110*(43), 17223–17228. doi:10.1073/pnas.1306973110

Seinfeld, J. H., & Pankow, J. F. (2003). Organic atmospheric particulate material. *Annual Review of Physical Chemistry*, *54*(1), 121–140. doi:10.1146/annurev.physchem.54.011002.103756

Shrivastava, M., Cappa, C. D., Fan, J., Goldstein, A. H., Guenther, A. B., Jimenez, J. L., ... Zhang, Q. (2017). Recent advances in understanding secondary organic aerosol: Implications for global climate forcing. *Reviews of Geophysics*, *55*(2), 509–559. doi:10.1002/2016RG000540

Simpson, D., Winiwarter, W., Börjesson, G., Cinderby, S., Ferreiro, A., Guenther, A., ... Öquist, M. G. (1999). Inventorying emissions from nature in Europe. *Journal of Geophysical Research: Atmospheres*, *104*(D7), 8113–8152. doi:10.1029/98jd02747

Singh, H. B., & Hanst, P. L. (1981). Peroxyacetyl nitrate (PAN) in the unpolluted atmosphere: An important reservoir for nitrogen oxides. *Geophysical Research Letters*, *8*(8), 941–944. doi:10.1029/gl008i008p00941

Sipilä, M., Sarnela, N., Jokinen, T., Henschel, H., Junninen, H., Kontkanen, J., ... O'Dowd, C. (2016). Molecular-scale evidence of aerosol particle formation via sequential addition of HIO_3. *Nature*, *537*(7621), 532–534. doi:10.1038/nature19314

Stark, H., Yatavelli, R. L., Thompson, S. L., Kimmel, J. R., Cubison, M. J., Chhabra, P. S., ... Jimenez, J. L. (2015). Methods to extract molecular and bulk chemical information from series of complex mass spectra with limited mass resolution. *International Journal of Mass Spectrometry*, *389*, 26–38. doi:10.1016/j.ijms.2015.08.011

Thornton, J. A., Kercher, J. P., Riedel, T. P., Wagner, N. L., Cozic, J., Holloway, J. S., ... Brown, S. S. (2010). A large atomic chlorine source inferred from mid-continental reactive nitrogen chemistry. *Nature*, *464*(7286), 271–274. doi:10.1038/nature08905

Tröstl, J., Chuang, W. K., Gordon, H., Heinritzi, M., Yan, C., Molteni, U., ... Baltensperger, U. (2016). The role of low-volatility organic compounds in initial particle growth in the atmosphere. *Nature*, *533*(7604), 527–531. doi:10.1038/nature18271

Tunved, P., Hansson, H., Kerminen, V., Ström, J., Dal Maso, M., Lihavainen, H., ... Kulmala, M. (2006). High natural aerosol loading over boreal forests. *Science*, *312*(5771), 261–263. doi:10.1126/science.1123052

Twomey, S. (1977). Influence of pollution on shortwave albedo of clouds. *Journal of the Atmospheric Sciences*, *34*(7), 1149–1152. doi:10.1175/1520-0469(1977)034⟨1149:TIOPOT⟩2.0.CO;2

Tyndall, J. (1868). On a new series of chemical reactions produced by light. *Proceedings of the Royal Society of London*, *17*, 92–102. doi:10.1098/rspl.1868.0012

Tyndall, J. (1869). On the blue colour of the sky, the polarization of skylight, and on the polarization of light by cloudy matter generally. *Proceedings of the Royal Society of London*, *17*, 223–233. doi:10.1098/rspl.1868.0033

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., & Jimenez, J. L. (2009). Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data. *Atmospheric Chemistry and Physics*, *9*(9), 2891–2918. doi:10.5194/acp-9-2891-2009

Valiev, R. R., Hasan, G., Salo, V.-T., Kubečka, J., & Kurten, T. (2019). Intersystem crossings drive atmospheric gas-phase dimer formation. *The Journal of Physical Chemistry A*, *123*(30), 6596–6604. doi:10.1021/acs.jpca.9b02559

Vereecken, L., & Francisco, J. (2012). Theoretical studies of atmospheric reaction mechanisms in the troposphere. *Chemical Society Reviews*, *41*(19), 6259–6293. doi:10.1039/c2cs35070j

Weber, R. J., & McMurry, P. H. (1996). Fine particle size distributions at the Mauna Loa Observatory, Hawaii. *Journal of Geophysical Research: Atmospheres*, *101*(D9), 14767–14775. doi:10.1029/95JD02271

Went, F. W. (1960). Blue hazes in the atmosphere. *Nature*, *187*(4738), 641–643. doi:10.1038/187641a0

Wimmer, D., Mazon, S. B., Manninen, H. E., Kangasluoma, J., Franchin, A., Nieminen, T., ... Petäjä, T. (2018). Ground-based observation of clusters and nucleation-mode particles in the Amazon. *Atmospheric Chemistry and Physics*, *18*(17), 13245–13264. doi:10.5194/acp-18-13245-2018

Yan, C., Nie, W., Äijälä, M., Rissanen, M. P., Canagaratna, M. R., Massoli, P., ... Ehn, M. (2016). Source characterization of highly oxidized multifunctional compounds in a boreal forest environment using positive matrix factorization. *Atmospheric Chemistry and Physics*, *16*(19), 12715–12731. doi:10.5194/acp-16-12715-2016

Yli-Juuti, T., Nieminen, T., Hirsikko, A., Aalto, P. P., Asmi, E., Hõrrak, U., ... Riipinen, I. (2011). Growth rates of nucleation mode particles in Hyytiälä during 2003-2009: Variation with particle size, season, data analysis method and

ambient conditions. *Atmospheric Chemistry and Physics*, *11*(24), 12865–12886. doi:10.5194/acp-11-12865-2011

Zeebe, R. E., Ridgwell, A., & Zachos, J. C. (2016). Anthropogenic carbon release rate unprecedented during the past 66 million years. *Nature Geoscience*, *9*(4), 325–329. doi:10.1038/ngeo2681

Zha, Q., Yan, C., Junninen, H., Riva, M., Sarnela, N., Aalto, J., ... Ehn, M. (2018). Vertical characterization of highly oxygenated molecules (HOMs) below and above a boreal forest canopy. *Atmospheric Chemistry and Physics*, *18*(23), 17437–17450. doi:10.5194/acp-18-17437-2018

Zhang, H., Yee, L. D., Lee, B. H., Curtis, M. P., Worton, D. R., Isaacman-VanWertz, G., ... Goldstein, A. H. (2018). Monoterpenes are the largest source of summertime organic aerosol in the southeastern United States. *Proceedings of the National Academy of Sciences*, *115*(9), 2038–2043. doi:10.1073/pnas.1717513115

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., ... Worsnop, D. R. (2007). Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. *Geophysical Research Letters*, *34*(13). L13801. doi:10.1029/2007GL029979

Ziemann, P. J., & Atkinson, R. (2012). Kinetics, products, and mechanisms of secondary organic aerosol formation. *Chemical Society Reviews*, *41*(19), 6582. doi:10.1039/c2cs35122f